ARTICLES

Cost of Hospitalization for Preterm and Low Birth Weight Infants in the United States
Rebecca B. Russell, Nancy S. Green, Claudia A. Steiner, Susan Meikle, Jennifer L. Howse, Karalee Poschman, Todd Dias, Lisa Potetz, Michael J. Davidoff, Karla Damus, and Joann R. Petrini

Magnetic Resonance Imaging Regional T1 Abnormalities at Term Accurately Predict Motor Outcome in Preterm Infants
Yukiko Nanba, Kiyoshi Matsui, Noriko Aida, Yoshiaki Sato, Katsuaki Toyoshima, Motoyoshi Kawataki, Rikuo Hoshino, Makiko Ohyama, Yasufumi Itani, Akiko Goto, and Akira Oka
Pediatrics 2007; 120: e10-e19.

Physician Medical Decision-making at the End of Life in Newborns: Insight Into Implementation at 2 Dutch Centers
A. A. Eduard Verhagen, Mark A. H. van der Hoeven, R. Corine van Meerveld, and Pieter J. J. Sauer

Hand-Washing and Diapering Equipment Reduces Disease Among Children in Out-of-Home Child Care Centers
Pediatrics 2007; 120: e29-e36.

Enzyme Replacement Therapy in Patients Who Have Mucopolysaccharidosis I and Are Younger Than 5 Years: Results of a Multinational Study of Recombinant Human α-L-Iduronidase (Laronidase)
J. Edmond Wraith, Michael Beck, Roderick Lane, Ans van der Ploeg, Elsa Shapiro, Yong Xue, Emil D. Kakkis, and Nathalie Guffon

Successful Intermittent Prophylaxis With Trimethoprim/Sulfamethoxazole 2 Days per Week for Pneumocystis carinii (jiroveci) Pneumonia in Pediatric Oncology Patients
Susan Lindemulder and Edythe Albano
Pediatrics 2007; 120: e47-e51.

Preventable and Unpreventable Causes of Childhood-Onset Epilepsy Plus Mental Retardation
Carol Camfield and Peter Camfield
Pediatrics 2007; 120: e52-e55.

Detection and Significance of Serum Protein Marker of Hirschsprung Disease
Jia-xiang Wang, Pan Qin, Qiu-liang Liu, He-ying Yang, Ying-zhong Fan, Jie-kai Yu, and Shu Zheng
Pediatrics 2007; 120: e56-e60.

Obesity Risk for Female Victims of Childhood Sexual Abuse: A Prospective Study
Jennie G. Noll, Meg H. Zeller, Penelope K. Trickett, and Frank W. Putnam

Evaluating Deliberation in Pediatric Primary Care
Elizabeth D. Cox, Maureen A. Smith, and Roger L. Brown
Pediatrics 2007; 120: e68-e77.

Prenatal Cocaine and Tobacco Effects on Children's Language Trajectories
Barbara A. Lewis, H. Lester Kirchner, Elizabeth J. Short, Sonia Minnes, Paul Weishampel, Sudtida Satayathum, and Lynn T. Singer
Pediatrics 2007; 120: e78-e85.

Self-Reported Health Status and Health-Related Quality of Life of Teenagers Who Were Born Before 29 Weeks' Gestational Age
Ron Gray, Stavros Petrou, Christine Hockley, and Frances Gardner
Pediatrics 2007; 120: e86-e93.
Patterns and Interrelationships of Body-Fat Measures Among Rural Chinese Children Aged 6 to 18 Years  
Hongjian Wang, Rachel E. Story, Scott A. Venners, Binyan Wang, Jianhua Yang, Zhiping Li, Liuliu Wang,  
Xue Liu, Genfu Tang, Houxun Xing, Xiping Xu, and Xiaobin Wang  

Growth and Growth Hormone Therapy in Subjects With Mulibrey Nanism  
Niklas Karlberg, Hannu Jalanko, and Marita Lipsanen-Nyman  
Pediatrics 2007; 120: e102-e111.

Parental Ability to Discriminate the Weight Status of Children: Results of a Survey  
Jeannie S. Huang, Karen Becerra, Thaghar Oda, Edward Walker, Ronghui Xu, Michael Donohue, Iris Chen,  
Violeta Curbelo, and Adam Breslow  
Pediatrics 2007; 120: e112-e119.

Hematologic and Blood Biochemistry Monitoring During Methylphenidate Treatment in Children With Attention-Deficit/Hyperactivity Disorder: 2-Year, Open-Label Study Results  
Sharon B. Wigal, Timothy E. Wilens, Mark Wolraich, and Marc Lerner  
Pediatrics 2007; 120: e120-e128.

Abnormal Thermoregulatory Responses in Adolescents With Chronic Fatigue Syndrome: Relation to Clinical Symptoms  
Vegard Bruun Wyller, Kristin Godang, Lars Mørkrid, Jerome Philip Saul, Erik Thaulow, and Lars Walløe  
Pediatrics 2007; 120: e129-e137.

Use of Complementary and Alternative Medicine in a General Pediatric Clinic  
Dany Jean and Claude Cyr  
Pediatrics 2007; 120: e138-e141.

Frequency and Characteristics of Pediatric and Adolescent Visits in Naturopathic Medical Practice  
Wendy Weber, James A. Taylor, Rachelle L. McCarty, and Amy Johnson-Grass  
Pediatrics 2007; 120: e142-e146.

Cardiovascular Effects of Sibutramine in the Treatment of Obese Adolescents: Results of a Randomized, Double-Blind, Placebo-Controlled Study  
Stephen R. Daniels, Barbara Long, Scott Crow, Dennis Styne, Melinda Sothern, Ileana Vargas-Rodriguez,  
Lisa Harris, Julia Walch, Olga Jasinsky, Kristine Cwik, Ann Hewkin, Vicky Blakesley for the Sibutramine Adolescent Study Group  

Motor Profile of Children With Developmental Speech and Language Disorders  
Chris Visscher, Suzanne Houwen, Erik J.A. Scherder, Ben Moolenaar, and Esther Hartman  
Pediatrics 2007; 120: e158-e163.

Prolonged Hypocortisolemia in Hydrocortisone Replacement Regimens in Adrenocorticotropic Hormone Deficiency  
Ann M. Maguire, Geoffrey R. Ambler, Bin Moore, Mark McLean, Marina G. Falleti, and Christopher T. Cowell  
Pediatrics 2007; 120: e164-e171.

Mass Distribution of Free, Intranasally Administered Influenza Vaccine in a Public School System  
L. Rand Carpenter, John Lott, Brian M. Lawson, Stephanie Hall, Allen S. Craig, William Schaffner, and  
Timothy F. Jones  

Rapid-Onset Obesity With Hypothalamic Dysfunction, Hypoventilation, and Autonomic Dysregulation Presenting in Childhood  
Diego Ize-Ludlow, Juliette A. Gray, Mark A. Sperling, Elizabeth M. Berry-Kravis, Jeff M. Milunsky, I. Sadaf  
Farooqi, Casey M. Rand, and Debra E. Weese-Mayer  
Pediatrics 2007; 120: e179-e188.

Human Milk Intake and Retinopathy of Prematurity in Extremely Low Birth Weight Infants  
Cherrie D. Heller, Michael O'Shea, Qing Yao, John Langer, Richard A. Ehrenkranz, Dale L. Phelps, W.  
Kenneth Poole, Barbara Stoll, Shahnaz Duara, William Oh, James Lemons, Brenda Poindexter for the NICHD Neonatal Research Network  
Malpractice Claims Involving Pediatricians: Epidemiology and Etiology
Aaron E. Carroll and Jennifer L. Buddenbaum
Pediatrics 2007; 120: 10-17.

Addressing Parents’ Concerns About Childhood Immunizations: A Tutorial for Primary Care Providers
Benjamin H. Levi

A Decline in the Frequency of Neonatal Exchange Transfusions and Its Effect on Exchange-Related Morbidity and Mortality
Laurie A. Steiner, Matthew J. Bizzarro, Richard A. Ehrenkranz, and Patrick G. Gallagher

Characteristics of the Pediatric Hospitalist Workforce: Its Roles and Work Environment
Gary L. Freed, Kathryn Brzoznowski, Kamilah Neighbors, Indu Lakhani American Board of Pediatrics, Research Advisory Committee

Growth and Neurodevelopmental Outcomes After Early Low-Dose Hydrocortisone Treatment in Extremely Low Birth Weight Infants

Botulinum Toxin for Spasticity in Children With Cerebral Palsy: A Comprehensive Evaluation
Kristie Bjornson, Ross Hays, Cathy Graubert, Robert Price, Francine Won, John F. McLaughlin, and Morty Cohen

Early Intervention and Recovery Among Children With Failure to Thrive: Follow-up at Age 8
Maureen M. Black, Howard Dubowitz, Ambika Krishnakumar, and Raymond H. Starr, Jr

Preterm Infants as Young Adults: A Swedish National Cohort Study
Karolina Lindström, Birger Winbladh, Bengt Haglund, and Anders Hjern
Pediatrics 2007; 120: 70-77.

Three-Tesla Cardiac Magnetic Resonance Imaging for Preterm Infants
Adrienne M. Foran, Julie A. Fitzpatrick, Joanna Allsop, Stephan Schmitz, Jamie Franklin, Constandinos Pamboucas, Declan O'Regan, Jo V. Hajnal, and A. David Edwards
Pediatrics 2007; 120: 78-83.

Hydrops Fetalis: A Retrospective Review of Cases Reported to a Large National Database and Identification of Risk Factors Associated With Death
Matthew E. Abrams, Keith S. Meredith, Paula Kinnard, and Reese H. Clark
Pediatrics 2007; 120: 84-89.

Cost-effectiveness of Alternative Strategies for Tuberculosis Screening Before Kindergarten Entry
Valerie J. Flaherman, Travis C. Porco, Elliot Marseille, and Sarah E. Royce

Steven R. Nesheim, Bill G. Kapogiannis, Minn M. Soe, Kevin M. Sullivan, Elaine Abrams, John Farley, Paul Palumbo, Linda J. Koenig, and Marc Bulterys
Pediatrics 2007; 120: 100-117.

Is Victimization From Bullying Associated With Medicine Use Among Adolescents? A Nationally Representative Cross-sectional Survey in Denmark
Pernille Due, Ebba Holme Hansen, Juan Merlo, Anette Andersen, and Bjørn E. Holstein
Pediatrics 2007; 120: 110-117.

Mental Health and Social Competencies of 10- to 12-Year-Old Children Born at 23 to 25 Weeks of Gestation in the 1990s: A Swedish National Prospective Follow-up Study
Aijaz Farooqi, Bruno Hägglöf, Gunnar Sedin, Leif Gothefors, and Fredrik Serenius
Pediatrics 2007; 120: 118-133.
Children Plus All Nonautomobile Motorized Vehicles (Not Just All-Terrain Vehicles) Equals Injuries
Christy L. Collins, Gary A. Smith, and R. Dawn Comstock
Pediatrics 2007; 120: 134-141.

Injuries in Canadian Youth Ice Hockey: The Influence of Relative Age
Nick Wattie, Stephen Cobley, Alison Macpherson, Andrew Howard, William J. Montelpare, and Joseph Baker

SPECIAL ARTICLES

Screening and Treatment for Lipid Disorders in Children and Adolescents: Systematic Evidence Review for the US Preventive Services Task Force
Elizabeth M. Haney, Laurie Hoyt Huffman, Christina Bougatsos, Michele Freeman, Robert D. Steiner, and Heidi D. Nelson
Pediatrics 2007; 120: e189-e214.

Screening for Lipid Disorders in Children: US Preventive Services Task Force Recommendation Statement
US Preventive Services Task Force

2006 Job Lewis Smith Award Acceptance Address: Is There an OWA (Other Weird Arrangement) in Your Future?
Stanley I. Fisch

COMMENTARIES

Determinants of Outcome After Head Cooling for Neonatal Encephalopathy
John M. Keogh
Pediatrics 2007; 120: 171-172.

Pediatricians and Medical Malpractice
Steven M. Donn and Gary McAbee

Bilirubin Toxicity to Human Erythrocytes: A More Sanguine View
Antony F. McDonagh

Hypnosis as a Therapeutic Tool in Pediatrics
Haleh Saadat and Zeev N. Kain

Counting Children With Disability in Low-Income Countries: Enhancing Prevention, Promoting Child Development, and Investing in Economic Well-being
Michael E. Msall and Dennis P. Hogan

Pay-for-Performance in Pediatrics: Proceed With Caution
Alyna T. Chien and R. Adams Dudley
Pediatrics 2007; 120: 186-188.
FROM THE AMERICAN ACADEMY OF PEDIATRICS

Hepatitis A Vaccine Recommendations
Committee on Infectious Diseases
Pediatrics 2007; 120: 189-199.

Preparation for Emergencies in the Offices of Pediatricians and Pediatric Primary Care Providers
Committee on Pediatric Emergency Medicine
Pediatrics 2007; 120: 200-212.

School Transportation Safety
Committee on Injury, Violence, and Poison Prevention and Council on School Health

Prevention of Varicella: Recommendations for Use of Varicella Vaccines in Children, Including a Recommendation for a Routine 2-Dose Varicella Immunization Schedule
Committee on Infectious Diseases
Pediatrics 2007; 120: 221-231.

EXPERIENCE & REASON

Mollaret Meningitis Associated With an Intraspinal Epidermoid Cyst
Baoqin Gao, Jun Yang, Shougang Zhuang, Yaxian Deng, Weili Yang, Yazhen Yu, Yongjun Wang, Lin Luo, and Ke Dai

LETTERS TO THE EDITOR

Stressing About Posttraumatic Stress Disorder
R. Grant Steen and Robert M. Hamer

Stressing About Posttraumatic Stress Disorder: In Reply
Victor G. Carrion, Carl F. Weems, and Allan L. Reiss

Increased Cerebral Blood Flow Velocity in Children With Sickle Cell Disease: Adenotonsillectomy or Transfusion Regimens?
Brigitte Bader-Meunier, Martine François, Suzanne Verlhac, Monique Elmaleh, Ghislaine Ithier, and Malika Benkerrou

Increased Cerebral Blood Flow Velocity in Children With Sickle Cell Disease: Adenotonsillectomy or Transfusion Regimens?: In Reply
Alexandra M. Hogan, Catherine M. Hill, Romola Bucks, Paul Telfer, and Fenella J. Kirkham

Do We Need to Incorporate Pharmacogenetics in Randomized, Controlled Trials of Frequently Used Medicines?
John N. van den Anker

Do We Need to Incorporate Pharmacogenetics in Randomized, Controlled Trials of Frequently Used Medicines?: In Reply
Amy C. Plint and Eric Clark
Self-injectable Epinephrine for First-Aid Management of Anaphylaxis
Alan D. Gaines
Pediatrics 2007; 120: 238.

Self-injectable Epinephrine for First-Aid Management of Anaphylaxis
Marcus Shaker, Donald Woodmansee, M. Kay Wolfson, and David Goodman
Pediatrics 2007; 120: 238-239.

Self-injectable Epinephrine for First-Aid Management of Anaphylaxis: In Reply
Scott H. Sicherer and F. Estelle R. Simons

Ventilatory Pump Failure and Strategies to Prevent Bronchopulmonary Dysplasia
Jon F. Watchko
Pediatrics 2007; 120: 240.

Ventilatory Pump Failure and Strategies to Prevent Bronchopulmonary Dysplasia: In Reply
Hany Aly

Newborn Screening for Cystic Fibrosis in New York State
Michele Caggana and Laura Helton

Newborn Screening for Cystic Fibrosis in New York State: In Reply
Robert Giusti, Ashley Badgwell, and Alejandro D. Iglesias

Prenatal Chlorpyrifos and Early Neurodevelopment: How Good Is the Science?
Domenic V. Cicchetti

Prenatal Chlorpyrifos and Early Neurodevelopment: How Good Is the Science?: In Reply
Pediatrics 2007; 120: 243-244.

Pediatric Palliative Care
Stacey Howell
Pediatrics 2007; 120: 244-245.

Pediatric Palliative Care: In Reply
Gina Santucci, Tammy I. Kang, and Chris Feudtner
Pediatrics 2007; 120: 245.

Choosing the Best Practice: Evidence to Support Fluoroquinolone Drops for Acute Otitis Media Through Tympanostomy Tubes
Joseph Dohar, William Giles, Peter Roland, Nadim Bikhazi, Sean Carroll, Roderick Moe, Bradley Reese, Sheryl Dupre, Michael Wall, David Stroman, Celeste McLean, Krista Crenshaw, Ramzi Younis, and Michael Poole

Choosing the Best Practice: Evidence to Support Fluoroquinolone Drops for Acute Otitis Media Through Tympanostomy Tubes: In Reply
Glenn Isaacson

Neonatal Blue-Light Phototherapy Could Increase the Risk of Dysplastic Nevus Development
Phyllis A. Dennery and Scott Lorch

Paucity of Evidence-Based Research on How to Achieve the Healthy People 2010 Goal of Exclusive Breastfeeding
Karen A. Bonuck

Much Pain, Little Gain From Voiding Cystourethrogram After Urinary Tract Infection
Eduardo Humberto Garin and Linda Young
ERRATA

ERRATA

ERRATA
Cost of Hospitalization for Preterm and Low Birth Weight Infants in the United States

Rebecca B. Russell, MSPH, Nancy S. Green, MD, Claudia A. Steiner, MD, MPH, Susan Meikle, MD, MPH, Jennifer L. Howse, PhD, Karalee Poschman, MPH, Todd Dias, MS, Lisa Potetz, MPP, Michael J. Davidoff, MPH, Karla Damus, PhD, RN, Joann R. Petrini, PhD, MPH

March of Dimes, White Plains, New York; Departments of Pediatrics and Obstetrics and Gynecology and Women’s Health, Albert Einstein College of Medicine, Bronx, New York; Center for Delivery, Organization and Markets, Agency for Healthcare Research and Quality, Rockville, Maryland

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

ABSTRACT

OBJECTIVE. The objective of this study was to estimate national hospital costs for infant admissions that are associated with preterm birth/low birth weight.

METHODS. Infant (<1 year) hospital discharge data, including delivery, transfers, and readmissions, were analyzed by using the 2001 Nationwide Inpatient Sample from the Healthcare Cost and Utilization Project. The Nationwide Inpatient Sample is a 20% sample of US hospitals weighted to approximately 35 million hospital discharges nationwide. Hospital costs, based on weighted cost-to-charge ratios, and lengths of stay were calculated for preterm/low birth weight infants, uncomplicated newborns, and all other infant hospitalizations and assessed by degree of prematurity, major complications, and expected payer.

RESULTS. In 2001, 8% (384,200) of all 4.6 million infant stays nationwide included a diagnosis of preterm birth/low birth weight. Costs for these preterm/low birth weight admissions totaled $5.8 billion, representing 47% of the costs for all infant hospitalizations and 27% for all pediatric stays. Preterm/low birth weight infant stays averaged $15,100, with a mean length of stay of 12.9 days versus $600 and 1.9 days for uncomplicated newborns. Costs were highest for extremely preterm infants (<28 weeks’ gestation/birth weight <1000 g), averaging $65,600, and for specific respiratory-related complications. However, two thirds of total hospitalization costs for preterm birth/low birth weight were for the substantial number of infants who were not extremely preterm. Of all preterm/low birth weight infant stays, 50% identified private/commercial insurance as the expected payer, and 42% designated Medicaid.

CONCLUSIONS. Costs per infant hospitalization were highest for extremely preterm infants, although the larger number of moderately preterm/low birth weight infants contributed more to the overall costs. Preterm/low birth weight infants in the United States account for half of infant hospitalization costs and one quarter of pediatric costs, suggesting that major infant and pediatric cost savings could be realized by preventing preterm birth.
Prematurity is the leading cause of neonatal mortality and a major cause of pediatric morbidity and disability, associated with up to one half of all pediatric neurodevelopmental disorders. Furthermore, preterm birth and low birth weight (LBW) may also be associated with lifelong chronic conditions, such as hypertension and dyslipidemia. The rates of preterm birth (<37 completed weeks of gestation) in the United States have been steadily increasing to a current level of 12.3% of all 4 million annual births, higher than the rates in most major industrialized countries. Compared with term births, infant mortality rates are 15-fold and 75-fold higher for those who are born preterm and very preterm (<32 weeks), respectively. Approximately 10% of all newborns are admitted to NICUs, many because of prematurity. Despite advancements in technologies and treatments in the past decade, the incidence of severe acute complications for very preterm/LBW infants, accompanied by risks for chronic medical conditions, have not markedly diminished since the mid-1990s.

A major proportion of pediatric hospital stays in the United States are for conditions in the neonatal period, which are among the most expensive diagnoses for all children. Previous studies have described a sizable financial toll from prematurity. Schmitt et al and Marbella et al both demonstrated in state-specific studies that preterm births contributed disproportionately to overall delivery costs, accounting for a small percentage of discharges and approximately half of all costs. Other studies have shown that hospital costs decrease with increasing birth weight and gestational age, with the smallest and earliest infants having the highest costs and longest length of stay (LOS). However, most of these studies focused on older data sets, population subsets such as newborns who were privately insured, employer-sponsored health plans, and data sets that were limited to local or regional reporting.

Marked shifts in treatment of high-risk mothers and infants, increasing viability for the earliest preterm infants, rapidly rising health care expenditures, and regional variations in preterm birth and NICU care have an impact on infant health care costs. We conducted a national analysis to estimate the US inpatient hospital costs for infants (<1 year of age) that are associated with preterm births and LBW, including deliveries and readmissions, using a nationally representative hospital discharge database.

METHODS

Study Design
A cross-sectional study of preterm birth/LBW-associated discharges, costs, and corresponding LOS for US infant hospitalizations was conducted using the 2001 Nationwide Inpatient Sample (NIS). The analysis included stratification by infant death during hospitalization, specific complications associated with preterm birth, and expected source of payment. The contribution of preterm/LBW infant hospitalization costs toward pediatric hospital costs was assessed.

Source of Data
The 2001 NIS database from the Healthcare Cost and Utilization Project (HCUP) is sponsored by the Agency for Healthcare Research and Quality in partnership with 33 state data organizations. The NIS sampling design approximates a 20% sample of US “community” hospitals from participating states. A community hospital is defined by the American Hospital Association as “all nonfederal, short-term, general, and other specialty hospitals, excluding hospital units of institutions,” which includes long-term hospitals, short-term rehabilitation hospitals, psychiatric hospitals, and alcoholism/chemical dependency treatment facilities. This hospital definition includes public, private, and academic medical centers, inclusive of those that provide maternity and newborn services. Hospitals were sampled on the basis of 5 strata (geographic region, rural/urban, number of beds, teaching status, and ownership), and all discharges from a sampled hospital are included in the NIS. Sample weights are provided and allow for analyses of nationwide estimates. Details of the NIS sample design and method for generating the weights to produce national estimates are described elsewhere. The 2001 NIS contains 7,452,727 unweighted inpatient hospital stays from a total of 986 US hospitals from states that participate in HCUP and after weighting approximates the 35 million inpatient hospital stays in the United States annually, making it the largest all-payer inpatient care database that is publicly available in the United States.

Variable Definitions
Infant hospital stays were defined as all admissions, including the newborn admission at delivery, hospital transfers, and all readmissions up to 1 year of age. The NIS contains data regarding charges; LOS; diagnoses and procedures using the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes; admission source; and transfer and discharge status. For newborn deliveries, the mother and the infant have separate admissions and are recorded as 2 individual stays in the NIS. Hospital stays were defined using the 15 primary and secondary ICD-9-CM diagnoses listed in the NIS and categorized into groups defined as preterm birth and/or LBW, uncomplicated newborn, and all other infants. Preterm/LBW infants were all infant stays with a diagnosis of short gestation, low birth weight, or slow fetal growth and malnutrition (ICD-9-CM codes 764–765 and V21.3). Most of the infants in this category were preterm and may or may not have been LBW. A small proportion of stays in the preterm/LBW group were assigned to this group because of an ICD-9-CM code of...
Source: Agency for Healthcare Research and Quality, 2001 NIS. Number of discharges and total costs for these categories exceed 100% because of overlapping diagnoses, but total costs for any diagnosis of preterm birth/LBW were not double counted. CI indicates
conﬁdence interval.

20 096–36 244
734–2024
1 115 676–1 330 477
4 521 906–6 154 936
28 170
1379
1 223 076
5 338 421
700
2600
500
800
1400–2500
3200–7900
600–700
2000–2700
1900
5600
600
2300
2
4
2
2
2.7–3.1
4.4–11.3
1.9–2.0
2.8–3.1
12 700–16 400
200–300
1 776 300–2 083 300
2 123 800–2 471 000
14 500
200
1 929 800
2 297 400

2.9
7.8
1.9
3.0

4 896 718–6 677 146
4 717 296–6 444 176
1 544 728–2 282 655
3 152 723–4 213 007
2 696 633–3 618 846
435 643–584 147
14 737–25 738
367 207–522 893
341 762–487 304

95% CI
In Thousands

5 786 932
5 580 736
1 913 692
3 682 865
3 157 740
509 895
20 237
445 050
414 533
3000
4000
36 800
3500
7500
1200
2300
1100
1500
13 300–16 800
14 700–18 600
57 700–73 500
10 800–13 400
14 900–18 300
3900–5100
7300–12 100
6500–8700
8100–10 800
15 100
16 700
65 600
12 100
16 600
4500
9700
7600
9500
4
5
26
5
9
3
3
3
3
12.0–13.7
13.1–14.9
38.5–45.9
10.7–12.0
14.6–16.1
4.4–5.0
6.4–9.3
7.1–8.5
8.6–10.3
12.9
14.0
42.2
11.4
15.3
4.7
7.9
7.8
9.4
352 200–416 100
305 700–363 400
25 200–33 100
279 800–331 100
173 000–207 900
103 700–122 100
1700–2400
53 300–63 900
39 900–47 800
384 200
334 500
29 200
305 500
190 500
112 900
2100
58 600
43 800

Total
n

95% CI

Mean

95% CI

Median

Mean

95% CI

Median

Cost
LOS
Discharges
Classiﬁcation

Preterm/LBW
Short gestation/LBW (ICD-9-CM codes 765.00–765.19)
Extreme Immaturity (ICD-9-CM codes 765.00–765.09)
Other preterm (ICD-9-CM codes 765.10–765.19)
⬍2500 g (ICD-9-CM codes 765.11–765.18)
ⱖ2500 g (ICD-9-CM code 765.19)
Birth weight not stated (ICD-9-CM code 765.10)
Slow growth/malnutrition (ICD-9-CM codes 764.00–764.99)
⬍2500 g (ICD-9-CM codes 764.01–764.08 764.11–764.18,
764.21–764.28, 764.91–764.98)
ⱖ2500 g (ICD-9-CM codes 764.09, 764.19, 764.29, 764.99)
Birth weight not stated (ICD-9-CM codes 764.00, 764.10, 764.20, 764.90)
Uncomplicated newborns
All other infants

Statistical Analyses
Data on hospital costs are presented as a better estimate
of economic impact than the corresponding charges,22,23,28 which reflect pricing decisions that take into
account payer polices and other factors that are unrelated to resource use.29 Hospital charges were used to
estimate costs by applying the cost-to-charge ratio (C/C)

TABLE 1 Hospital Costs and LOS for Infant Hospitalizations: United States, 2001

slow fetal growth/malnutrition, and these stays may also
have been LBW. All preterm/LBW hospitalizations were
then identified as LBW (⬍2500 g), very low birth weight
(⬍1500 g), not LBW (ⱖ2500 g), or birth weight not
stated using the fifth digit of the ICD-9-CM codes.
Uncomplicated newborn stays were defined as all
stays at delivery with a principal diagnosis of liveborn
(ICD-9-CM codes V30.0 –V39.2, excluding stillbirths,
V32, V35, or V36, and any diagnosis of disease). All
other infant stays included stays that did not meet the
criteria for the previous 2 groups and included admissions at delivery and readmissions with a diagnosis of
disease other than preterm birth/LBW.
Preterm/LBW stays were stratified into short gestation/LBW or slow growth/fetal malnutrition (see Table 1
for specific ICD-9-CM codes) for better understanding
the contribution of each of the groups to the overall costs
for preterm/LBW infant stays. Because of the 15 diagnoses included in the NIS, it was possible for 1 stay to
include a code for both of these groups, making them
not mutually exclusive; however, this occurred in ⬍1%
of stays. Within the short gestation/LBW group, analyses
were further separated to distinguish stays with a diagnosis of “extreme immaturity” or extremely preterm/
LBW (gestation of ⬍28 completed weeks and/or birth
weight of ⬍1000 g) and “other preterm” (gestation of
28 –36 completed weeks of gestation and/or birth weight
of 1000 –2499 g), with a very small amount of overlap
between these 2 groups (⬍0.01%). Slow growth/malnutrition and other preterm stays were further stratified
using the detail provided by the fifth digit of the ICD9-CM codes that reflect birth weight, which were collapsed into the categories ⬍2500 g, ⱖ2500 g, and birth
weight not stated.
Survival to discharge or death before discharge was
determined using the discharge status reported in the
NIS. Complications of preterm birth studied included
respiratory distress syndrome (RDS), bronchopulmonary dysplasia (BPD), intraventricular hemorrhage
(IVH), and necrotizing enterocolitis (NEC; ICD-9-CM
codes shown in Table 2). Mechanical ventilation was
defined using ICD-9 procedure codes 93.90 to 93.91,
96.04, and 96.70 to 96.72.
Source of payment reflects expected payer at time of
discharge. An algorithm using primary and secondary
payer variables was used to assign expected payer to 1 of
4 categories: Medicaid, private insurance, uninsured (including self-pay), and other (including Medicare).

PEDIATRICS Volume 120, Number 1, July 2007

e3


TABLE 2  Selected Complications for Preterm/LBW Infants Who Weighed <2500 g Among Infants Who Were Not Transferred and Infants Who Were Transferred From Another Acute Care Hospital: United States, 2001

<table>
<thead>
<tr>
<th>Condition/Procedure</th>
<th>Discharges</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>95% CI</td>
</tr>
<tr>
<td>Death</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Died during stay</td>
<td>11 200</td>
<td>9600–12 800</td>
</tr>
<tr>
<td>Survived to discharge</td>
<td>202 600</td>
<td>184 400–220 800</td>
</tr>
<tr>
<td>RDS (ICD-9-CM code 769)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition present</td>
<td>49 900</td>
<td>42 600–57 100</td>
</tr>
<tr>
<td>Condition absent</td>
<td>164 100</td>
<td>150 700–177 500</td>
</tr>
<tr>
<td>BPD (ICD-9-CM code 770.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition present</td>
<td>9400</td>
<td>7100–11 700</td>
</tr>
<tr>
<td>Condition absent</td>
<td>204 600</td>
<td>186 800–223 300</td>
</tr>
<tr>
<td>IVH (ICD-9-CM codes 772.10–772.14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition present</td>
<td>8900</td>
<td>7200–10 700</td>
</tr>
<tr>
<td>Condition absent</td>
<td>205 000</td>
<td>186 800–223 300</td>
</tr>
<tr>
<td>NEC (ICD-9-CM code 777.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition present</td>
<td>2900</td>
<td>2300–3600</td>
</tr>
<tr>
<td>Condition absent</td>
<td>211 000</td>
<td>191 900–230 100</td>
</tr>
</tbody>
</table>

Source: Agency for Healthcare Research and Quality, 2001 NIS.

developed by Agency for Healthcare Research and Quality using data from the Centers for Medicare and Medicaid Services for nearly every hospital that participated in HCUP in 2001.30,31 The 2001 group weighted average C/C for each hospital in the NIS is a weighted average of C/C for all hospitals in a group, which is determined by state, urban/rural, investor owned/other, and number of beds.30 For stays that were missing data on charges (2.8% of infant stays) or missing a C/C (15.8% of infant stays), cost was imputed by assigning the mean cost for that analytic group. Stays with imputed costs were then included in total cost estimates.

Although it was not possible to track hospital stays for the same patient when transferred or readmitted because the NIS discharge records lack identifying information, algorithms were established using NIS variables to minimize the potential small error from overlapping hospital stays for transferred patients. The discharge disposition, admission source, and the ICD-9-CM codes for delivery from the discharge record were used to stratify the infant hospitalizations into 4 categories for analysis: not transferred, transferred out, transferred in, and readmitted (see Table 3 for details).

Potential double counting of transferred or readmitted infants was minimized by examination of deaths and complications only among 2 mutually exclusive groups of preterm/LBW infants: deliveries not transferred and infants transferred in from another hospital. The analysis was then further limited to the higher risk group of LBW stays within this group. The mean and median costs that were associated with death during hospitalization and specific complications were calculated.

Analyses of costs and LOS exclude outliers, defined as stays with charges less than $25 and greater than or equal to $1 000 000 and LOS >365 days, which were 0.03% of all discharges in 2001. SAS 8.2 (SAS Institute, Cary, NC) was used to conduct statistical analyses. All estimates of number of discharges and costs were rounded to the nearest hundred. PROC SURVEYMEANS in SAS was used to calculate appropriate SEs and 95% confidence intervals for discharges, mean LOS, and mean and total costs, taking into account the sampling design of the NIS.32 Significant differences in mean costs between 2 independent samples were assessed using a t test calculated in SAS. Differences in proportions of payer coverage between preterm/LBW infants and uncomplicated newborns were assessed using a χ² test of proportions using EpiInfo 3.3.2 (Centers for Disease Control and Prevention, Atlanta, GA).

RESULTS

Infant Hospitalizations

Costs for all 4.6 million infant hospitalizations totaled an estimated $12.4 billion (Fig 1). Eight percent (384 200) of the total 4.6 million infant stays included a diagnosis of preterm birth/LBW yet accounted for 47% of all infant costs ($5.8 billion; Fig 1). Sixty-six percent (253 200) of all preterm/LBW infant hospitalizations had an ICD-9-CM code that indicated that they were LBW and 17% were very LBW (<1500 g) and averaged $20 600 and $52 300, respectively. LBW discharges accounted for 90% ($5.2 billion) of the total costs. In comparison, uncomplicated newborns composed 42% of infant stays (1 929 800) but only 10% of all infant costs ($1.2 billion). The remaining infant hospitalizations accounted for 50% of all stays and 43% of all costs. These remaining infants represent a wide range of conditions and diagnoses, encompassing >2100 individual ICD-9-CM codes identifying the principal reasons for the stay.
and a total of >4300 different ICD-9-CM codes listed as the secondary diagnosis, with an interquartile range of costs from $500 to $1600 (25th and 75th percentiles) and a maximum cost of nearly $546,000.

Infant hospitalization costs totaled 59% of all pediatric (<18 years) hospital costs, which totaled $21.2 billion for the 6.4 million admissions in 2001. Hospitalization costs for preterm/LBW infants composed 27% of costs for all pediatric stays.

The mean cost for preterm/LBW infant stays was $15,100, compared with $600 for an uncomplicated newborn and $2300 for all other infant hospitalizations (Table 1). The mean LOS for preterm/LBW, uncomplicated, and all other infants were 12.9, 1.9, and 3.0 days, respectively (Table 1). Median cost and LOS for preterm/LBW infants were lower than the mean cost and LOS, indicating a subset of costly stays that reflect an intense use of resources among more complicated stays.

### Categories of Preterm/LBW Infants

Costs were estimated for preterm/LBW infants who were diagnosed as short gestation and/or LBW compared with those who were growth restricted or coded as small for gestational age. More than 87% of preterm/LBW infants (334,500) were of short gestation/LBW (Table 1), whereas 15% (58,600) had slow growth/malnutrition, which included small for gestational age. The majority of stays with a diagnosis of slow growth/malnutrition were <2500 g (43,800). Costs for short gestation/LBW totaled $5.6 billion and nearly $450 million for slow growth/malnutrition.

Hospitalizations for extremely preterm/LBW infants (<28 weeks’ gestation/birth weight <1000 g) represented 8% of preterm/LBW infants and had a mean cost of $65,600 and mean LOS of 42.2 days. Other preterm infants (28–36 weeks’ gestation/birth weight 1000–2499 g) had a mean cost of $12,100 and an average LOS of 11.4 days. The majority of these other preterm stays had a birth weight of <2500 g. Average costs for these infants were nearly 4 times higher than for other preterm infants with a birth weight of ≥2500 g: $16,600 and $4500, respectively.

### Admission and Transfer Status

Of all hospitalizations for preterm/LBW infants, 80% (306,000) were deliveries that involved no hospital transfer (Table 3). These infants had a mean LOS of 12 days and mean costs of nearly $13,000. Eight percent of preterm/LBW infant admissions (30,000) were deliveries that resulted in a transfer to another acute care hospital, and half of these occurred on the day of delivery. Seven percent (26,100) of preterm/LBW stays were for infants who were admitted from another acute care hospital. The remaining 6% of all preterm/LBW infant stays (22,100) were readmissions after hospital discharge after delivery.
Infants who were transferred in from another hospital had the highest mean cost ($40 800) and the longest mean LOS (26.6 days) of all admission categories ($P < .001). On average, readmissions had a LOS of 16.4 days and a mean cost of $16 400. Similar trends were seen for transfers and readmissions when the analysis was limited to preterm infants with a code that indicated LBW (data not shown).

**Infant Complications**

In 2001, 5.2% of preterm/LBW infants who weighed <2500 g died before discharge, compared with 3.5% of preterm infants of all birth weights (data not shown). Mean hospital costs for surviving infants were not significantly different for those who died before discharge ($21 400 and $20 400, respectively; $P = .65). However, the median cost for deaths was $2100, whereas the median cost among survivors was $7300.

Mean hospital costs for preterm/LBW infants who weighed <2500 g and had specific complications were 4 to 7 times higher than those for infants without these complications. One fourth of stays studied had 1 or more of the following complications: RDS, BPD, IVH, and NEC. Costs for these 4 conditions totaled $3.1 billion. Among these preterm/LBW infants, the single costliest complication in terms of average cost per discharge was BPD, with an average cost of $116 000, reported in 4.4% of cases (Table 2). The most common complication was RDS, reported in 23.3%, or 49 900, infants. Average cost for stays with RDS was $56 800, compared with $10 700 for infants without RDS. IVH was reported in 4.2% and NEC in 1.4% of all preterm/LBW infants who weighed <2500 g (Table 2). These infants had high hospital costs: $76 000 for IVH and $100 000 for NEC. Use of mechanical ventilation was identified among 27.3% of preterm/LBW infants who weighed <2500 g, with costs averaging $55 100.

**Source of Expected Payer**

Among preterm/LBW infant stays, 42% had Medicaid listed as either the primary or the secondary expected payer, 50% had private/commercial insurance, 5% listed uninsured/self-pay, and 2% listed other sources as the expected payer, such as Medicare (Table 4). The mean costs for stays were similar whether covered by Medicaid ($15 800) or by private/commercial insurance ($15 000; $P = .58). These costs were approximately twice as high as uninsured/self-pay stays ($8700; $P < .001). In contrast, regardless of payer group, costs for uncomplicated newborn stays averaged $600. When

<table>
<thead>
<tr>
<th>Expected Payer</th>
<th>Preterm/LBW Infants</th>
<th>Uncomplicated Newborns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (95% CI)</td>
<td>%</td>
</tr>
<tr>
<td>Medicaid</td>
<td>160 700 (144 200–177 200)</td>
<td>41.9</td>
</tr>
<tr>
<td>Private/commercial</td>
<td>193 300 (171 700–214 900)</td>
<td>50.4</td>
</tr>
<tr>
<td>Uninsured/self-pay</td>
<td>20 300 (14 700–25 900)</td>
<td>5.3</td>
</tr>
<tr>
<td>Other*</td>
<td>9000 (6700–11 400)</td>
<td>2.3</td>
</tr>
</tbody>
</table>

Source: Agency for Healthcare Research and Quality, 2001 NIS.  
* Includes Medicare.
compared with preterm/LBW stays, a significantly lower proportion of uncomplicated stays were covered by Medicaid (37.5%; P < .001) than by private/commercial insurance (54.5%; P < .001). The proportions of those designated as uninsured/self-pay (5.5%) and other sources (2.5%) were slightly larger among uncomplicated newborns (P < .001).

DISCUSSION
Preterm birth is a major obstetric and pediatric challenge, because it is a common, persistent, and often devastating condition with substantial medical, economic, emotional, and social impact. Our study, which includes all expected payers, quantifies on a nationwide scale the vastly disproportionate infant hospitalization costs that are associated with preterm birth/LBW: 47% of all infant costs for only 8% of all infant admissions, as well as 27% of all pediatric hospital costs for 6% of all pediatric admissions. Mean hospital costs for preterm/LBW infants were 25 times higher and average hospitalizations were 11 days longer than those for uncomplicated births. Marked disparate expenses for preterm infants include intensive care and treatment for acute and chronic conditions. As expected, respiratory complications were the most prevalent and costly for preterm/LBW infants who require transfer in the first few days of life. Therefore, continued focused efforts to address causes for very preterm and moderate to late preterm births remain imperative. In addition, the subset of preterm infants who require transfer in the first few days of life require the greatest intervention.

Second, our results are consistent with those from previous studies that demonstrated that the initial (ie, delivery) hospital stay incurred the majority of hospital expenses during the first year of life for preterm/LBW infants. Readmissions during the first year of life were only a small proportion of these preterm/LBW infant stays, accounting for only 6% of hospital admissions and $363 million, or 6%, of total costs. In contrast, the 7% of infants who were transferred in from another hospital incurred by far the highest mean costs and LOS. Therefore, continued focused efforts to address causes for very preterm and moderate to late preterm births remain imperative. Therefore, continued focused efforts to address causes for very preterm and moderate to late preterm births remain imperative.

Finally, private and public health care purchasers share the financial burden and, hence, the impact of preterm birth. It is interesting that Medicaid covered nearly 42% of preterm/LBW infant hospitalizations, compared with nearly 38% of uncomplicated newborn stays. This reflects the demographics of populations with higher preterm birth rates but may also include infants who qualify for Medicaid as a result of the costliness of care that is required for preterm infants.

Although the use of discharge data has coding limitations, this study used strengths of the NIS, including the detailed ICD-9-CM codes for birth weight, to target certain analyses to high-risk infants, as well as admission source and disposition to account for transfers. Analysis of transfers suggests that a small number of all preterm/LBW stays included in this study were transfers and thus may potentially lead to double counting of infants. Importantly, the identified number of infants whose discharge disposition indicated transfer to another facility nearly matched the identified number of infants whose admission source indicated a transfer in from another facility, a finding that supports internal validity of the data.

We acknowledge the limitations that are inherent in this study of hospitalization costs using discharge data. There is agreement within the literature that both hospital charge and cost estimates have limitations and that precise costs are difficult to derive because of the inherent limitations of standard cost conversion methods. Specifically, C/Cs used here do not recognize differences in markups among services within a hospital, which may affect the estimated costs for preterm infants relative to others. The NIS database uses ICD-9-CM codes to capture the diagnoses on the hospital discharge record for all hospitalizations. Because the actual gestational age of each infant is not available in the 2001 database, we may be underestimating a portion of total number of discharges and total costs for preterm/LBW infants if the stays did not receive the proper ICD-9-CM codes or if the prematurity was not clinically significant.

Total health care costs for preterm/LBW infants are stabilized in the past several decades, at nearly matched the identified number of infants whose discharge disposition indicated transfer to another facility nearly matched the identified number of infants whose admission source indicated a transfer in from another facility, a finding that supports internal validity of the data.

We acknowledge the limitations that are inherent in this study of hospitalization costs using discharge data. There is agreement within the literature that both hospital charge and cost estimates have limitations and that precise costs are difficult to derive because of the inherent limitations of standard cost conversion methods. Specifically, C/Cs used here do not recognize differences in markups among services within a hospital, which may affect the estimated costs for preterm infants relative to others. The NIS database uses ICD-9-CM codes to capture the diagnoses on the hospital discharge record for all hospitalizations. Because the actual gestational age of each infant is not available in the 2001 database, we may be underestimating a portion of total number of discharges and total costs for preterm/LBW infants if the stays did not receive the proper ICD-9-CM codes or if the prematurity was not clinically significant.

Total health care costs for preterm/LBW infants are

Total health care costs for preterm/LBW infants are...
certainly higher than our estimated costs, because inpatient costs in the NIS do not include physician and other professional fees, rehabilitation, or outpatient or home care expenses. Additional data sets linking hospital discharge with other databases, such as vital statistics and those that include outpatient care, could provide a more detailed picture of these costs but may be possible only by restricting analysis to a single health care system,\(^{18,20,39}\) precluding generalization to national health care costs. Maternal costs that are associated with hospitalization and treatment of preterm labor, which have been shown to be substantially greater than costs for term deliveries,\(^{36}\) are another major expense of preterm birth. Other maternal costs include outpatient treatment, employer health care expenses, and lost work productivity. A recently released Institute of Medicine report\(^{36}\) on preterm birth estimated the total societal economic cost of preterm birth to be $26 billion, including medical costs from birth through early childhood, maternal delivery costs, early intervention and special education services, and lost household and labor market productivity. Although annual national inpatient hospital costs for preterm birth, extrapolated to the United States from a cohort of 24,000 Utah births that were covered by a single health care plan, were approximately twice as high as those found in our analyses, methodologic differences between these 2 estimates can likely explain the disparity.

The causes of preterm labor and birth are not wholly understood, and few successful interventions have been determined. Causes of preterm births at \(\geq 32\) weeks’ gestation may be more amenable to prevention\(^{34}\) than extremely preterm births and therefore may be a potential source of future cost savings and improvements in developmental outcomes. Promising prevention strategies include infection prevention, smoking cessation, and 17α-hydroxyprogesterone therapy to prevent recurrence among eligible women.\(^{36}\) Our data on costs for specific neonatal morbidities that are associated with preterm birth also suggest that interventions that are aimed at prevention of complications could lead to significant cost savings.

**CONCLUSIONS**

This study provides a robust national estimate of infant hospitalization costs that are associated with preterm birth/LBW and demonstrates that health care that is associated with these outcomes compose a substantial portion of the cost of hospital care for all infants and children, especially for those who are at the lowest gestational ages. As the incidence of preterm birth grows\(^{4}\) and the limits of viability extend,\(^{1}\) the economic impact of preterm birth will likely grow. Ongoing trend analyses by the authors indicate a rise in the hospital price tag for preterm/LBW infants (data not shown). The enormous costs of treatment for preterm infants relative to total infant and pediatric costs highlight the need for enhanced efforts and research that are targeted at prevention of preterm birth\(^{34}\) and that would benefit private and public payers alike.

**ACKNOWLEDGMENTS**

We gratefully acknowledge the expert input from John M. Tillford, PhD (Department of Pediatrics, University of Arkansas for Medical Science); Norman Waitzman, PhD (University of Utah); and Jeannette Rogowski, PhD (University of Medicine and Dentistry of New Jersey). We also acknowledge the statewide data organizations that participated in the HCUP NIS 2001: Arizona Department of Health Services; California Office of Statewide Health Planning & Development; Colorado Health & Hospital Association; Connecticut-CHIME, Inc; Florida Agency for Health Care Administration; Georgia-GHA: An Association of Hospitals & Health Systems; Hawaii Health Information Corporation; Illinois Health Care Cost Containment Council; Iowa Hospital Association; Kansas Hospital Association; Kentucky Department for Public Health; Maine Health Data Organization; Maryland Health Services Cost Review Commission; Massachusetts Division of Health Care Finance and Policy; Michigan Health and Hospital Association; Minnesota Hospital Association; Missouri Hospital Industry Data Institute; Nebraska Hospital Association; New Jersey Department of Health & Senior Services; New York State Department of Health; North Carolina Department of Health and Human Services; Oregon Association of Hospitals & Health Systems; Pennsylvania Health Care Cost Containment Council; Rhode Island Department of Health; South Carolina State Budget & Control Board; Tennessee Hospital Association; Texas Health Care Information Council; Utah Department of Health; Vermont Association of Hospitals and Health Systems; Virginia Health Information; Washington State Department of Health; West Virginia Health Care Authority; and Wisconsin Department of Health & Family Services.

**REFERENCES**

and risk of cardiovascular disease in a cohort of women followed up since 1976. BMJ. 1997;315:396–400
Magnetic Resonance Imaging Regional T1 Abnormalities at Term Accurately Predict Motor Outcome in Preterm Infants

Yukiko Nanba, MDa,b, Kiyoshi Matsui, MDa, Noriko Aida, MDb, Yoshiaki Sato, MDa, Katsuaki Toyoshima, MDa, Motoyoshi Kawataki, MDa, Rikuo Hoshino, MDa, Makiko Ohyama, MDb, Yasufumi Itani, MDb, Akiko Goto, MDb, Akira Oka, MDb

Departments of aNeonatology and bRadiology, Kanagawa Children's Medical Center, Yokohama, Japan; cDivision of Perinatology, Department of Perinatal Medicine and Maternal Care, and dDivision of Child Neurology, Department of Medical Subspecialties, National Center for Child Health and Development, Tokyo, Japan

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

ABSTRACT

OBJECTIVE. The aim of this study was to assess whether periventricular leukomalacia findings are sufficiently sensitive for predicting the severity of motor prognosis by conventional MRI in the near term.

METHODS. Preterm infants with T1 hyperintensity or cysts in the periventricular regions on term MRI were selected, and their gross motor functions were evaluated at the age of 3 to 5 years. Sixty-two infants had findings of T1 hyperintensity or cysts, and except for infants with these findings, none were diagnosed later as periventricular leukomalacia.

RESULTS. All 37 patients with cerebral palsy had periventricular lesions with T1 hyperintensity or cysts in the corona radiata above the posterior limb of the internal capsule on coronal sections. Small T1 hyperintensity lesions were seen on coronal slices and were often difficult to detect on axial slices. All of the 17 infants with T1 hyperintensity findings sparing the corona radiata above the posterior limb of the internal capsule showed normal motor development, irrespective of findings of ventriculomegaly. There was a tendency for the presence of widespread lesions in corona radiata above the posterior limb of the internal capsule to be correlated with the severity of motor handicap.

CONCLUSIONS. Lesions in the corona radiata above the posterior limb of the internal capsule on a coronal view by term MRI were useful for predicting motor prognosis in preterm infants with periventricular leukomalacia.
PERIVENTRICULAR LEUKOMALACIA (PVL) is a major type of brain injury in preterm infants. Ultrasonography and MRI are the standard methods that are used for diagnosing PVL. Ultrasonography clearly detects cystic PVL at the bedside, and MRI is superior to ultrasonography in detecting noncystic white matter lesions.1-5

A number of studies have demonstrated correlations between motor prognosis and PVL findings by cranial MRI.6 These MRI scans were conducted serially over periods of months or several years (late MRI). PVL on late MRI was diagnosed when there were (1) ventriculomegaly with an irregular outline of the body and trigone of the lateral ventricles, (2) a reduced quantity of periventricular white matter, and (3) abnormal signal intensity in the periventricular white matter.7,8

Several studies have reported that neonatal conventional MRI, at around term (term MRI), could be used to predict cerebral palsy (CP) with a high degree of sensitivity and specificity.5,9–12 The PVL findings in these studies were cystic lesions, diffuse T2 hyperintensity, and T1 hyperintensity. In a previous report by us, as well as others, T1 hyperintensity lesions in the periventricular white matter were shown to be associated with PVL on late MRI.7,8,12-15 Some infants with these lesions have developed normally without CP,12,16 and a correlation of PVL findings in term MRI with the severity of motor problems has not been examined. The accurate identification of PVL before discharge is clinically important for the early prediction of motor sequelae and for targeting high-risk infants to appropriate rehabilitation services. The aim of this study was to assess whether the findings of T1 hyperintensity in periventricular white matter by conventional MRI around term are correlated with PVL. Term MRI was usually performed before discharge from the hospital using a 1.5T scanner and consisted of the following: coronal and axial spin-echo (SE) T1-weighted images (400/15/2 [echo time/repetition time/excitation]) and coronal and/or axial SE T2-weighted images (300/81–120/1) with 5-mm slices. All MRIs were evaluated independently by 3 of the authors (Drs. Nanba, Matsui, and Aida) without knowledge of the clinical outcome, and in case of discrepancy, the findings were established after the discussion by 3. We selected infants with MRI findings of T1 shortening or cysts in the periventricular white matter, which had been reported as findings of PVL at term MRI.7,8 The following MRI findings were studied: (1) distribution of T1 hyperintensity, (2) the presence or distribution of cystic lesions, (3) degree of ventriculomegaly, and (4) degree of irregularity of the ventricular outline. Ventriculomegaly was divided into 5 groups: none, slight, mild when the ventricular/brain ratio (V/B) at the level of the midbody of the lateral ventricles was <0.34 but the occipital horn was largely dilated, moderate when the V/B exceeded 0.35, and severe when little white matter was seen because of the dilation of ventricles.9 The irregularity of the ventricular outline was classified into 4 degrees: none, slight, apparent, and severe.

Intraventricular hemorrhage (IVH) showed a T1 hyperintensity in the periventricular region on term MRI. IVHs were observed to be of homogeneous extreme hyperintensity in the T1 sequence with T2 hypointensity along the ventricular margin. IVH often coexisted with PVL, but to clarify the prognosis for the infants with the lesion in the periventricular white matter, we excluded patients with IVH, including germinal matrix hemorrhage.

Myelination of the corona radiata (CR) also showed T1 hyperintensity in the periventricular region. The normal
myelination of CR has been reported to be evident on axial slices after the corrected age of 36 weeks.8,18,19 In terms of the normal myelination of corticospinal tracts in the periventricular region projecting from the motor cortex down to the posterior limb of the internal capsule (PLIC), which was easily observed in coronal images, T1 hyperintensity appeared after the corrected age of 44 weeks (Fig 2). Therefore, it was difficult to distinguish PVL lesions in periventricular CR related to the corticospinal tract (CR-CSp) from normal myelination, and MRIs that were taken

FIGURE 1
Numbers of infants who were studied.

2342 infants were admitted between January 1993 and July 2000
1119 infants were born at a gestational age of 24–34 wk
460 MRI scans: all infants with a birth weight of <1500 g or abnormal findings
430 MRIs were carried out at corrected age of 36-43 weeks
13 infants died after discharge
→ 26 were not followed because of hospital transfer or removal
102 infants were excluded for the following reasons:
   Intracranial hemorrhages: 61 cases
   Hydrocephalus: 23 cases
   Brain malformations: 3 cases
   Chromosomal abnormalities: 2 cases
   Anomaly syndrome: 7 cases
   Congenital myotonic dystrophy: 3 cases
   Congenital viral infections: 3 cases
289 infants were eligible
   62 had a lesion in the periventricular white matter
     45 infants with the lesion in the CR-CSp
       37 infants were spastic palsy due to PVL
       8 had normal neuromotor development
     17 infants with the lesion sparing the CR-CSp
     227 had no lesion in term MRI
     226 infants had normal neuromotor development
     1 infant had hypotonic palsy without the findings of PVL in late MRI

FIGURE 1
Numbers of infants who were studied.
between postmenstrual weeks 36 and 43 were included in this study.

Of the 62 infants studied, follow-up MRIs were performed 1 to 3 years later for 17 infants for clinical indications. Late MRI included axial SE T1-weighted images (360–500/15/2) and T2-weighted images (3000/80–110/1).

All MRIs were performed with the infants in stable condition, and the infants were sedated with pentobarbital (2.5–10 mg/kg body weight). Heart rate and transcutaneous oxygen saturation were monitored during and after the examination, and all examinations were done in a safe manner. Informed consent was obtained from all parents, and the study was based on the ethical guidelines approved by the ethics committee of the hospital.

**Gross Motor Function**

Infants with the lesions in the periventricular white matter and abnormal neurologic examination were followed up for at least 3 years by pediatricians who had been trained in these procedures. The severity of CP was classified into levels I to V on the basis of the Gross Motor Function Classification System (GMFCS)\(^2\) (Table 1). No abnormality in gross motor development was classified into level 0.

<table>
<thead>
<tr>
<th>Level</th>
<th>Gross Motor Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Walk without restrictions; limitations in more advance motor skills</td>
</tr>
<tr>
<td>II</td>
<td>Walk without assistive devices; limitations walking outdoors and in the community</td>
</tr>
<tr>
<td>III</td>
<td>Walk with assistive mobility devices; limitations walking outdoors and in the community</td>
</tr>
<tr>
<td>IV</td>
<td>Self-mobility with limitations; transported or uses power mobility outdoors and in the community</td>
</tr>
<tr>
<td>V</td>
<td>Self-mobility is severely limited even with assistive technology</td>
</tr>
</tbody>
</table>
RESULTS
Of 289 eligible infants, 62 infants had findings of T1 hyperintensity or cystic lesions in the periventricular white matter and were followed for 3 to 5 years. Of 227 infants with normal term MRI findings, none later received a diagnosis as PVL. It was difficult to identify any particular antecedents from the NICU course, which might be responsible for each MRI finding.

Characteristics of T1 Hyperintensity Lesions
Distributed T1 hyperintensity lesions were observed on both coronal and axial images on term MRI in 9 patients. These were sometimes linear heterogeneous punctate lesions in the periventricular white matter in axial images. Where spotty T1 hyperintensity lesions were seen, the T2 image showed mild hypointensity or isointensity (Figs 3 and 4). Small T1 hyperintensity lesions were seen as spotty on coronal slices in 35 patients. These were often difficult to find on axial slices because of the T1 hyperintensity of normal myelination (Fig 3). These small punctate T1 hyperintensity lesions showed an isointensity on T2 images. As a result, it was essential to investigate coronal T1 images, to detect all of the lesions in the periventricular white matter. Twenty-four of 62 patients had a cystic PVL on term MRI. In the cases with cystic PVL, T1 hyperintensity lesions were often in proximity to regions of cystic changes (Fig 4).

Relation Between T1 Hyperintensity or Cystic Lesions and GMFCS
Of 62 infants with the T1 hyperintensity, 25 had normal motor development and 37 had CP. We prepared a list of these term-MRI findings that were classified by GMFCS level (Tables 2 and 3). All of the 17 infants with T1

---

FIGURE 3
A and B, Patient I-3. A, On a coronal view, spotty T1 hyperintensity lesions (arrowheads) were found in the CR-CSp. Linear myelination in the PLIC was seen (arrows). B, Lesions of T1 hyperintensity on coronal slices were not detected on axial slices. Axial T1-weighted image showed normal myelination in the CR (white arrows). C and D, Patient II-5. C, On a coronal slice, lesions of T1 hyperintensity in the CR-CSp (arrowheads) were also detected above the myelinated PLIC (black arrows). D, Cystic formation was evident (white arrowhead), but T1 hyperintensity lesions were not detected on axial slices. E-G, Patient IV-4. E, Coronal T1-weighted images showed band-like hyperintensity (arrowheads) above the PLIC (arrows). F and G, Cystic formation was evident (white arrowhead), and T2 hypo-or isointensity (G; arrowheads) lesions in the periventricular white matter. Slight ventriculomegaly and irregularity of ventricular outline were observed.
hyperintensity findings that spared the CR-CSp showed normal motor development without any signs of spasticity, irrespective of findings of ventriculomegaly or an irregularity in the ventricular outline (Table 2, Fig 5). In 4 infants with spotty T1 hyperintensity findings that spared the CR-CSp on term MRI, the late MRIs after the age of 1 showed widespread T2 hyperintensity in the periventricular white matter with a mild irregularity in the ventricular outline with or without a reduced quantity of periventricular white matter.

Among 45 children who showed T1 hyperintensity or cysts in the CR-CSp, 8 had normal development without any signs of spasticity and 37 had spastic motor defects as a result of PVL (Table 3). All 8 infants with normal motor development had small spotty T1 hyperintensity lesions, and none had widespread lesions or cysts in the CR-CSp; on their MRIs, a slight ventriculomegaly was sometimes seen, although irregularity of the ventricular walls was absent. Infants with findings in the CR-CSp, ventriculomegaly, and irregularity of the walls developed CP.

Sensitivity and Specificity of PVL Findings on Term MRI in Predicting the Severity of Motor Problems Among 289 Eligible Infants

The sensitivity and the specificity of cystic lesions in the periventricular white matter for detecting CP (GMFCS level I or higher) among the 289 infants were 62% (23 of 37) and 87% (251 of 289), respectively. The sensitivity and the specificity of lesions in the CR-CSp for detecting CP (GMFCS level of I or higher) among the 289 infants were 100% (37 of 37) and 97% (244 of 252), respectively (Table 4).

The widespread findings in the CR-CSp tended to correlate with a worse ventriculomegaly or irregularity in the ventricular wall. Of 13 infants with T1 hyperintensity in the CR-CSp and without any irregularity in

---

**FIGURE 4**

A–C, Patient V-7. A, The findings of T1 hyperintensity (arrowheads) and cysts (white arrowheads) in the CR-CSp proceeding to PLIC (black arrows) were seen on coronal images. T1 hyperintensity lesions were in proximity to a region of cystic change. B and C, Linear heterogeneous T1 hyperintensity and T2 hypo- or isointensity lesions (arrowheads) were seen on axial slices. The degree of ventriculomegaly was moderate and a severe irregularity in the ventricular outline was seen. D and E, Patient V-11. D, Severe cystic lesions (white arrowheads) were spread in the CR-CSp above myelinated PLIC (black arrows). D and E, There was no T1 hyperintensity lesion, and only cystic lesions (white arrowheads) were seen. Severe ventriculomegaly and irregularity in the ventricular wall were observed.
the ventricular outline, 8 had no abnormality in gross motor development (GMFCS 0) and 10 were able to walk without assistive devices (GMFCS 0–I). The lesions in the CR-CSp with ventriculomegaly (V/B H11350 0.35) were correlated with GMFCS V. These sensitivity and specificity for detecting GMFCS V were 100% (11 of 11) and 100% (278 of 278), respectively. Ventriculomegaly (V/B H11350 0.35) was always accompanied by an apparent irregularity in the ventricular outline. Lesions in the CR-CSp with an apparent irregularity in the ventricular wall were correlated with GMFCS IV or V. For these lesions, the sensitivity and the specificity for detecting GMFCS IV to V were 90% (18 of 20) and 100% (268 of 269), respectively.

There was a tendency for the severity of the MRI findings in CR-CSp to be correlated with the degree of motor disability. For detecting GMFCS levels I to III (they were able to walk with or without assistive devices), findings of T1 spotty hyperintensity without cysts in the CR-CSp had a sensitivity of 78% (14 of 18) and a specificity of 96% (260 of 269). For detecting GMFCS level IV (they were able to sit, but independent mobility was very limited), findings of T1 band-like hyperintensity in the CR-CSp had a sensitivity of 75% (6 of 8) and a specificity of 99% (278 of 281). The findings of cystic formation in the CR-CSp for detecting GMFCS V (they were not able to maintain antigravity head and trunk postures in prone and sitting positions; they were almost bedridden) had a sensitivity of 73% (8 of 11) and a specificity of 99% (275 of 278).

**DISCUSSION**

The major sequela of PVL is spastic diplegia, because most PVL occurs in the region of the white matter that is traversed by descending fibers from the motor cortex corresponding to the legs.6 We demonstrated that the presence of T1 hyperintensity lesions or cysts and their distribution at CR-CSp on a coronal view were important for detecting PVL related to motor defects. Coronal sections were superior to axial sections for detecting findings of T1 hyperintensity. On axial slices, we were unable to distinguish small spotty T1 hyperintensity lesions from myelination in the CR; as a result, punctate lesions on coronal slices were not often detected on axial views. On coronal slices, myelination in the CR above the PLIC that we noticed was not remarkable until a corrected age of 43 weeks, and we could easily define the lesions using the location of the PLIC as a hallmark. In previous studies, T1 hyperintensity was observed in preterm infants on MRI that was performed between the neonate and term-equivalent period.1–3,7,8,12–16,21 However, some reported that T1 hyperintensity did not affect the prognosis.16,21 These reports were based on axial or sagittal sections, and it might be difficult to determine the precise position of the findings with relation to the corticospinal tracts. We conclude that coronal MRIs between 36 and 43 weeks’ corrected age would be the most useful for diagnosis of PVL and the prediction of the severity of gross motor function. In this study, we evaluated MRIs with 5-mm slices; 2- to 3-mm slices would improve the sensitivity. Several studies of PVL investigated correlations between the severity of clinical features and findings in the corticospinal tract on late MRI but not on term MRI.22–25 In our experience, the distribution of T1 hyperintensity on term MRI was more limited than that of T2 or fluid-attenuated inversion-recovery hyperintensity on late MRI, and term MRI might more clearly demonstrate the focal lesions that are responsible for symptoms, rather than late MRI. Recent diffusion-weighted MRI studies

**TABLE 2 Term-MRI Findings: Patients With No Abnormality in the CR-CSp**

<table>
<thead>
<tr>
<th>Patient</th>
<th>GA, wk</th>
<th>Birth Weight, g</th>
<th>CA at Examination, wk</th>
<th>Location of T1 Hyperintensity</th>
<th>Ventriculomegaly</th>
<th>Irregularity in the Ventricular Outline</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24</td>
<td>840</td>
<td>38</td>
<td>Anterior to the CR-CSp</td>
<td>Slight</td>
<td>None</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>915</td>
<td>38</td>
<td>Anterior to the CR-CSp</td>
<td>Slight</td>
<td>None</td>
</tr>
<tr>
<td>3</td>
<td>26</td>
<td>960</td>
<td>41</td>
<td>Anterior to the CR-CSp</td>
<td>Slight</td>
<td>None</td>
</tr>
<tr>
<td>4</td>
<td>26</td>
<td>890</td>
<td>41</td>
<td>Anterior to the CR-CSp</td>
<td>Slight</td>
<td>None</td>
</tr>
<tr>
<td>5</td>
<td>28</td>
<td>1160</td>
<td>38</td>
<td>Anterior to the CR-CSp</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>6</td>
<td>29</td>
<td>1190</td>
<td>42</td>
<td>Anterior to the CR-CSp</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>7</td>
<td>31</td>
<td>1510</td>
<td>36</td>
<td>Anterior to the CR-CSp</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>8</td>
<td>32</td>
<td>1222</td>
<td>38</td>
<td>Anterior to the CR-CSp</td>
<td>Mild</td>
<td>Slight</td>
</tr>
<tr>
<td>9</td>
<td>26</td>
<td>1000</td>
<td>41</td>
<td>Anterior and posterior to the CR-CSp</td>
<td>Slight</td>
<td>Slight</td>
</tr>
<tr>
<td>10</td>
<td>27</td>
<td>1018</td>
<td>41</td>
<td>Posterior to the CR-CSp</td>
<td>Mild</td>
<td>Slight</td>
</tr>
<tr>
<td>11</td>
<td>25</td>
<td>745</td>
<td>42</td>
<td>Posterior to the CR-CSp</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>12</td>
<td>27</td>
<td>885</td>
<td>41</td>
<td>Posterior to the CR-CSp</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>13</td>
<td>27</td>
<td>1176</td>
<td>40</td>
<td>Posterior to the CR-CSp</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>14</td>
<td>27</td>
<td>790</td>
<td>43</td>
<td>Posterior to the CR-CSp</td>
<td>Mild</td>
<td>Slight</td>
</tr>
<tr>
<td>15</td>
<td>31</td>
<td>1735</td>
<td>43</td>
<td>Posterior to the CR-CSp</td>
<td>Mild</td>
<td>Slight</td>
</tr>
<tr>
<td>16</td>
<td>32</td>
<td>1975</td>
<td>39</td>
<td>Posterior to the CR-CSp</td>
<td>Slight</td>
<td>Slight</td>
</tr>
<tr>
<td>17</td>
<td>33</td>
<td>1890</td>
<td>37</td>
<td>Posterior to the CR-CSp</td>
<td>Slight</td>
<td>Slight</td>
</tr>
</tbody>
</table>

GA indicates gestational age; CA, corrected age.
that were performed in the neonatal period revealed defect in the corticospinal tract with PVL. They may be used for detecting the focus of white matter abnormalities in the future, but this technique is difficult to use as a screening method.

The findings of T1 hyperintensity on term MRI pathologically corresponded to cellular reactions of glial cells and macrophages, as well as formation of microcalcifications. T1 hyperintensity lesions were often adjacent to regions of cystic change, consistent with necrotic changes without cyst formation. Because of the shrinkage of the necrotic lesions, the broad findings of T1 hyperintensity lesions as well as cystic lesions led to ventriculomegaly and an irregularity. The severe findings in CR-CSp associated with severe ventriculomegaly and irregularity resulted in the worst outcome.

The possibility that infants with T1 hyperintensity sparing CR-CSp have cognitive or other defects by reason of a reduction in white matter cannot be excluded. In this study, cognitive or visual dysfunctions were not evaluated, and additional study in this area is needed.

We examined the findings of T1 hyperintensity in preterm infants. Similar findings are also observed in term

<table>
<thead>
<tr>
<th>GMFCS</th>
<th>Patient</th>
<th>GA, wk</th>
<th>Birth Weight, g</th>
<th>CA at Examination, wk</th>
<th>Finding in the CR-CSp</th>
<th>Ventriculomegaly</th>
<th>Irregularity in the Ventricular Outline</th>
<th>Cyst Sparing the CR-CSp</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>26</td>
<td>815</td>
<td>42</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>2</td>
<td>27</td>
<td>1118</td>
<td>40</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>3</td>
<td>28</td>
<td>965</td>
<td>41</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>4</td>
<td>29</td>
<td>1170</td>
<td>38</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>5</td>
<td>31</td>
<td>1996</td>
<td>37</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>None</td>
<td>Apparent</td>
<td>None</td>
</tr>
<tr>
<td>6</td>
<td>26</td>
<td>930</td>
<td>37</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>Slight</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>7</td>
<td>31</td>
<td>1430</td>
<td>37</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>8</td>
<td>31</td>
<td>1430</td>
<td>39</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>I</td>
<td>1</td>
<td>24</td>
<td>715</td>
<td>39</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>750</td>
<td>41</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>Slight</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>1564</td>
<td>39</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>Slight</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>4</td>
<td>25</td>
<td>750</td>
<td>41</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>Slight</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>5</td>
<td>26</td>
<td>860</td>
<td>43</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>Mild</td>
<td>None</td>
<td>Slight</td>
</tr>
<tr>
<td>6</td>
<td>30</td>
<td>1530</td>
<td>40</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>Mild</td>
<td>Slight</td>
<td>None</td>
</tr>
<tr>
<td>II</td>
<td>1</td>
<td>29</td>
<td>1140</td>
<td>39</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>2</td>
<td>29</td>
<td>1440</td>
<td>38</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>Slight</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>3</td>
<td>29</td>
<td>1286</td>
<td>38</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>Slight</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>4</td>
<td>28</td>
<td>1100</td>
<td>38</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>Slight</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>1637</td>
<td>37</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>Slight</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>6</td>
<td>26</td>
<td>772</td>
<td>41</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>Slight</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>7</td>
<td>30</td>
<td>1466</td>
<td>40</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>Slight</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>III</td>
<td>1</td>
<td>31</td>
<td>1452</td>
<td>41</td>
<td>Spotty T1 hyperintensity</td>
<td>Slight</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>2</td>
<td>28</td>
<td>1452</td>
<td>38</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>Mild</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>830</td>
<td>40</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>Mild</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>4</td>
<td>32</td>
<td>2085</td>
<td>36</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>Mild</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>5</td>
<td>29</td>
<td>1386</td>
<td>37</td>
<td>Band-like hyperintensity</td>
<td>None</td>
<td>Mild</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>IV</td>
<td>1</td>
<td>30</td>
<td>1770</td>
<td>36</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>2</td>
<td>31</td>
<td>1640</td>
<td>39</td>
<td>Spotty T1 hyperintensity</td>
<td>None</td>
<td>Mild</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>3</td>
<td>28</td>
<td>1200</td>
<td>39</td>
<td>Band-like hyperintensity</td>
<td>None</td>
<td>Mild</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>4</td>
<td>29</td>
<td>1110</td>
<td>40</td>
<td>Band-like hyperintensity</td>
<td>None</td>
<td>Mild</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>5</td>
<td>28</td>
<td>1260</td>
<td>38</td>
<td>Band-like hyperintensity</td>
<td>None</td>
<td>Mild</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>6</td>
<td>27</td>
<td>950</td>
<td>40</td>
<td>Band-like hyperintensity</td>
<td>None</td>
<td>Mild</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>7</td>
<td>30</td>
<td>1288</td>
<td>37</td>
<td>Band-like hyperintensity</td>
<td>None</td>
<td>Mild</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>8</td>
<td>27</td>
<td>1020</td>
<td>39</td>
<td>Band-like hyperintensity</td>
<td>None</td>
<td>Mild</td>
<td>None</td>
<td>Apparent</td>
</tr>
<tr>
<td>V</td>
<td>1</td>
<td>27</td>
<td>1085</td>
<td>41</td>
<td>Spotty T1 hyperintensity</td>
<td>Moderate</td>
<td>Severe</td>
<td>Apparent</td>
</tr>
<tr>
<td>2</td>
<td>28</td>
<td>1050</td>
<td>40</td>
<td>Band-like hyperintensity</td>
<td>Moderate</td>
<td>Severe</td>
<td>Apparent</td>
<td>Apparent</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>1058</td>
<td>39</td>
<td>Band-like hyperintensity</td>
<td>Moderate</td>
<td>Severe</td>
<td>Apparent</td>
<td>Apparent</td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td>1624</td>
<td>38</td>
<td>Spotty T1 hyperintensity</td>
<td>Moderate</td>
<td>Severe</td>
<td>Apparent</td>
<td>Apparent</td>
</tr>
<tr>
<td>5</td>
<td>27</td>
<td>1175</td>
<td>38</td>
<td>Spotty T1 hyperintensity</td>
<td>Moderate</td>
<td>Severe</td>
<td>Apparent</td>
<td>Apparent</td>
</tr>
<tr>
<td>6</td>
<td>26</td>
<td>860</td>
<td>39</td>
<td>Spotty T1 hyperintensity</td>
<td>Moderate</td>
<td>Severe</td>
<td>Apparent</td>
<td>Apparent</td>
</tr>
<tr>
<td>7</td>
<td>28</td>
<td>1280</td>
<td>40</td>
<td>Spotty T1 hyperintensity</td>
<td>Moderate</td>
<td>Severe</td>
<td>Apparent</td>
<td>Apparent</td>
</tr>
<tr>
<td>8</td>
<td>26</td>
<td>905</td>
<td>40</td>
<td>Spotty T1 hyperintensity</td>
<td>Moderate</td>
<td>Severe</td>
<td>Apparent</td>
<td>Apparent</td>
</tr>
<tr>
<td>9</td>
<td>29</td>
<td>100</td>
<td>36</td>
<td>Spotty T1 hyperintensity</td>
<td>Moderate</td>
<td>Severe</td>
<td>Apparent</td>
<td>Apparent</td>
</tr>
<tr>
<td>10</td>
<td>34</td>
<td>1468</td>
<td>40</td>
<td>Spotty T1 hyperintensity</td>
<td>Moderate</td>
<td>Severe</td>
<td>Apparent</td>
<td>Apparent</td>
</tr>
<tr>
<td>11</td>
<td>26</td>
<td>1350</td>
<td>41</td>
<td>Spotty T1 hyperintensity</td>
<td>Moderate</td>
<td>Severe</td>
<td>Apparent</td>
<td>Apparent</td>
</tr>
</tbody>
</table>
infants, and it needs to be elucidated whether those findings in term infants have the same pathology as white matter injury in prematurity.

CONCLUSIONS

PVL lesions in term MRI cannot be correctly evaluated by axial slices, which are misleading. We conclude that coronal T1 sequences in term MRI are useful as a screening method for the diagnosis of PVL and the prediction of motor outcome in preterm infants.

ACKNOWLEDGMENTS

We thank Dr Yoshiaki Saito and Kousaku Ohno, Division of Child Neurology, Institute of Neurologic Sci-
ences, Tottori University School of Medicine, for help in the preparation of this manuscript.

REFERENCES


14. Keeney SE, Adcock EW, McArdle CB. Prospective observations of 100 high-risk neonates by high-field (1.5Tesa) magnetic resonance imaging of the central nervous system. II. Lesions associated with hypoxic-ischemic encephalopathy. Pediatrics. 1991;87:431–438


Physician Medical Decision-making at the End of Life in Newborns: Insight Into Implementation at 2 Dutch Centers

A. A. Eduard Verhagen, MD, JD*, Mark A. H. van der Hoeven, MD, PhDb, R. Corine van Meerveld, MDa, Pieter J. J. Sauer, MD, PhDa

aDepartment of Pediatrics, University Medical Center Groningen, Groningen, Netherlands; bDivision of Neonatology, University of Maastricht, Maastricht, Netherlands

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

ABSTRACT

OBJECTIVE. Decisions regarding end-of-life care in critically ill newborns in the Netherlands have received considerable criticism from the media and the public. This might be because of a lack of proper information and knowledge. Our purpose was to provide detailed information about how and when the implementation of end-of-life decisions, which are based on quality-of-life considerations, takes place.

METHODS. We reviewed the charts of all infants who died within the first 2 months of life at 2 university hospitals in the Netherlands from January to July 2005 and extracted all relevant information about the end-of-life decisions. We interviewed the responsible neonatologists about the end-of-life decisions and the underlying quality-of-life considerations and about the process of implementation.

RESULTS. Of a total of 30 deaths, 28 were attributable to withholding or withdrawing life-sustaining treatment. In 18 of 28 cases, the infant had no chance to survive; in 10 cases, the final decision was based on the poor prognosis of the infant. In 6 patients, 2 successive different end-of-life decisions were made. The arguments that most frequently were used to conclude that quality of life was deemed poor were predicted suffering and predicted inability of verbal and nonverbal communication. Implementation consisted of discontinuation of ventilatory support and alleviation of pain and symptoms. Neuromuscular blockers were added shortly before death in 5 cases to prevent gasping, mostly on parental request.

CONCLUSIONS. The majority of deaths were attributable to withholding or withdrawing treatment. In most cases, the newborn had no chance to survive and prolonging of treatment could not be justified. In the remaining cases, withholding or withdrawing treatment was based on quality-of-life considerations, mostly the predicted suffering and predicted inability of verbal and nonverbal communication. Potentially life-shortening medication played a minor role as a cause of death.
In the past 2 years, the foreign press (especially in Italy, the United Kingdom, and the United States) has paid extensive attention to a supposedly outrageous practice in the Netherlands of physicians’ terminating the life of severely defective newborn infants. It was suggested that all end-of-life (EoL) decisions in the Netherlands were in fact acts of euthanasia based on quality-of-life considerations, with reference to the medical practice in Germany during the Second World War. These accounts were based on the (mis)interpretation of publications regarding EoL decisions and the use of potentially life-shortening drugs in newborns in the Netherlands.

Decisions on when to start, withhold, or withdraw treatment in very sick newborns are among the most difficult decisions in pediatric practice. The difficulty lies to a large extent in the dilemma that is presented by 2 views: the value of life and the quality of life. The advances in technology and pharmacology have created new possibilities to save and prolong the life of newborns, but extension of life can also result in endless and severe suffering, which might not be in the interest of the infant. Studies from the United States and Europe have reported that the proportion of sick newborns in whom the decisions to withhold or withdraw life support preceded death increased substantially during the past 10 years. Quality-of-life concerns have been reported by neonatologists as reasons for these decisions in a substantial proportion (20%–50%) of deaths. Despite the frequency of these decisions, not much is known about what these quality-of-life concerns really are or about how the implementation of the decision takes place in practice. With respect to the latter, the role of potentially life-shortening drugs that alleviate pain and other symptoms at the end of life of newborns is of special interest. The legal difference between letting die and active ending of life is in principle based on the administration of these drugs.

The quality of EoL decisions, including the decisions regarding palliation of pain and symptoms, can be evaluated and compared between various countries only when sufficient insight is provided about medical practice at the end of life. With the purpose to gain more insight into what the medical practice in the Netherlands actually is, we conducted a retrospective descriptive study in 2 large university-based tertiary NICUs to determine the reasons that motivate physicians to make EoL decisions and how those are implemented in practice.

METHODS

Demographics

We reviewed the charts of all newborn infants who died in 2 university hospitals (A and B) with tertiary NICUs within the first 2 months of life between January and July 2005. We abstracted information from the attending physicians’ and nurses’ notes to determine demographics: birth weight, gestational age, day of death and diagnoses (using both clinical data and autopsy materials when available), and details about the decision-making process. According to Dutch law, no approval for this study from the ethical committee is required because it is a retrospective study using anonymous data.

The total number of NICUs in the Netherlands is limited to 10 by law to promote efficient use of expertise, manpower, and resources. All deliveries before 32 weeks’ gestation take place in a hospital with a NICU. Older newborn infants who require intensive care treatment are referred to these NICUs. Extracorporeal membrane oxygenation is not available in the study hospitals, and heart surgery is not available in NICU B. Patients who require these facilities are transferred to another NICU. In both NICUs, all medical decisions are made by a multidisciplinary team that is led by the attending neonatologist, who is ultimately responsible. The attending neonatologist also informs the parents of their infant’s status and proposed treatment plans during regular discussions with the parents. Discussions regarding limiting treatment options are initiated by both the parents and the physicians. Consensus among all team members and the parents in EoL decisions is always sought. During the study period, 423 patients were admitted (280 in NICU A and 143 in NICU B), and the average daily census of critically ill patients was 19 in NICU A and 13 in NICU B.

Classification of Newborns

The attending neonatologist’s daily notes and death summaries were used to determine whether death had occurred with or without a preceding medical EoL decision. Medical EoL decisions were defined as medical decisions with the effect or the probable effect that death was hastened. These decisions include the decisions to withhold or withdraw life-prolonging treatment and the decision to end deliberately the life of a newborn. On the basis of the notes, we categorized the newborns at the time of each EoL decision into 1 of the following groups from the literature: group 1, no chance to survive (NCTS); group 2, theoretical chance to survive, very poor prognosis (PP); or group 3, stable, hopeless prognosis with severe suffering, not depending on intensive care. During the study period, in neither center was a patient found to belong in group 3.

Decision-making and Implementation

Physicians’ notes were also reviewed to determine the physician’s reasons to withhold or withdraw life-prolonging treatment in all deaths. We also ascertained which individuals were involved in the decision-making process. The treatment orders that were given by the
physician on the basis of the EoL decision were collected from the files to describe the practice of implementation of these decisions. We used the medical charts and pharmacy notes to identify potentially life-shortening medication, comparing medication before and after the EoL decision. Medication before the decision was defined as the highest dosage of medication with potentially life-shortening effect as administered in the 12 hours before the EoL decision.

We interviewed all neonatologists who were involved in each EoL decision with quality-of-life arguments face to face. We cross-checked all data that were extracted from the medical charts and asked them to explain in detail why they took the decision to withhold or withdraw treatment and how the decision was implemented. In cases in which potentially life-shortening drugs were used, we asked for the purpose. The reasons to limit treatment were grouped into categories that were derived from publications in the Dutch medicolegal literature.22,23

Definitions
Withholding treatment was defined as withholding potentially life-saving treatment, which included not only withholding cardiopulmonary resuscitation but also not providing additional intensive care treatment (eg, not making additional ventilator changes despite hypoxemia, not providing additional catecholamines despite hypotension) in accordance with definitions in the 1992 report by the Dutch Pediatric Association.22 Withdrawing treatment was taken to be equivalent to withdrawing life-sustaining treatment (eg, withdrawing the ventilator). Deliberate ending of life was defined as administering lethal drugs with the purpose to end the life or shorten the life of a newborn who is otherwise stable. We do not use the term “euthanasia” because in the Netherlands, this can be used only when a physician ends the life of a patient on the patient’s explicit request, in accordance with the Dutch euthanasia law. In this legal framework, life shortening as an inevitable adverse effect of appropriate pain and/or symptom alleviation is considered acceptable clinical practice. The legal and moral status of administering lethal drugs with the purpose to shorten life in an unstable newborn in the dying phase, as part of careful EoL management, is still uncertain and subject to ongoing debate.

RESULTS
Demographics
A total of 30 newborns died within the first 2 months of life in the 2 hospitals during the 6-month study period: 21 died in hospital A, and 9 died in hospital B. Twenty-nine infants died in a NICU, and 1 patient died on a PICU. Table 1 shows the main characteristics of all deaths and the categories in which all newborns were classified. Overall, 24 (83%) deaths were attributable to withdrawal of treatment, 4 (10%) were by withholding treatment, and 2 (7%) occurred despite maximum treatment. The diagnoses that led to death varied, the largest group being term newborns with hypoxic-ischemic encephalopathy (23%).

Classification of Newborns
The data from the medical charts were sufficient to classify all patients in categories. Of all 28 deaths that were preceded by an EoL decision, 18 (64%) were classified as NCTS and 10 (36%) as PP at the time of the final decision. The proportion of deaths in the PP category was the same in both hospitals: 6 (35%) of 17 in hospital A and 4 (36%) of 11 in hospital B. In 9 (32%) of cases, 2 EoL decisions were made, with a median of 24 hours (range: 6–210 hours) between the first and the second decision. Six of these patients were initially classified as PP and moved to NCTS at the time of the second EoL decision. The proportion of deaths that were preceded by an EoL decision and classified as PP was the same in both hospitals: 6 (35%) of 17 in hospital A and 4 (36%) of 11 in hospital B.

Decision-making and Implementation
In all deaths in the NCTS group, it was apparent from the physicians’ notes that treatment was withdrawn because there was no chance of survival. Both the parents and the medical team consented to the decision in all documented cases (Table 2). The median time between the final decision and implementation of the decision was 2 hours (range: 0–24 hours). In 1 case, the time course was undocumented. The implementation consisted of discontinuation of artificial ventilation and removal of the endotracheal tube in all cases except 1. In 1 case, the patient was gradually weaned from the ventilator in 24 hours on parental request. Intravenous medication and fluids were continued until death in all cases. The median time between implementation and death was 30 minutes (range: 1–105 minutes).

In all cases of PP, treatment was withheld or withdrawn because the prognosis was considered very poor. The considerations that led to this conclusion were documented in the medical charts but without much detail (eg, treatment stopped because quality of life is deemed low). Table 3 shows more detailed information from the interviews indicating that in all cases, >1 consideration was present. The predicted very low quality of life was most frequently based on the predicted suffering and predicted inability to communicate. With inability to communicate, the neonatologists meant inability to be engaged in any kind of communication with other people, verbally or nonverbally (eg, because of deafness and blindness combined with predicted severe mental retardation or predicted vegetative state).

The median time between the decision and imple-
mentation of the decision in cases of PP was 1 hour (range: 0–24 hours). The implementation consisted of discontinuation of artificial ventilation and extubation in all cases except 1. In 1 case, the ventilatory support was weaned stepwise in 24 hours on the parents’ request. The median time between implementation and death in this group was 60 minutes (range: 15–360 minutes).

Each of the documented decisions was preceded by at least 1 or more decision-making meetings of the medical team followed by 1 or more with the parents. In these meetings, provision of sedation and analgesia as potentially life-shortening medication was also discussed. Table 4 presents the use and dosing of this medication before and after the final EoL decision. Before the final EoL decision was made, the majority of cases in NCTS (n = 15; 83%) and in PP (n = 9; 90%) received opioids and benzodiazepines. In all cases, the dosage was within the normal dosing range. Two newborns with pulmonary hypertension were treated with neuromuscular blockers (NMBs) as part of the hospital’s standard treatment of this disease. Additional medication after the final EoL decision was administered in 7 (39%) cases in NCTS and 8 (80%) cases in PP. It was given as an increased dosage of the existing continuous medication or as a bolus infusion. The dosages remained within normal dosing range. The reasons to provide additional medication were treatment of presenting symptoms (eg, pain, dyspnea, discomort) and prevention of suffering from these symptoms in the process of dying. None of the physicians interviewed considered hastening death as the aim of additional medication, but all declared that they would consider it an acceptable adverse effect. NMBs were added in 5 newborns in PP. Four of them had a diagnosis of hypoxic-ischemic encephalopathy, 1 with sepsis/meningitis. In 4 of 5 cases, NMBs were prescribed to prevent gasping. In 3 of these cases, this was done on explicit parental request, and in 1 case, it was the physician’s decision to do so because it was expected that gasping would scare the parents away from their dying child. In the remaining case, the dosage of previously prescribed NMB was increased to ensure optimal effect, whereas discontinuation was expected to impose unnecessary suffering. All cases were classified as deaths from natural cause by the attending neonatologists.
TABLE 2  Decision-making and Implementation in 30 Deaths

<table>
<thead>
<tr>
<th>Patient</th>
<th>Parental Consent</th>
<th>Team Decision</th>
<th>Time Between Last EoL Decision and Implementation, h:min</th>
<th>Time Between Implementation and Death, h:min</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>3</td>
<td>+</td>
<td>+</td>
<td>3.15</td>
<td>0.05</td>
</tr>
<tr>
<td>4</td>
<td>+</td>
<td>+</td>
<td>5.30</td>
<td>0.45</td>
</tr>
<tr>
<td>5</td>
<td>+</td>
<td>ND</td>
<td>4.30</td>
<td>0.30</td>
</tr>
<tr>
<td>6</td>
<td>+</td>
<td>+</td>
<td>1.45</td>
<td>0.30</td>
</tr>
<tr>
<td>7</td>
<td>ND</td>
<td>+</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>8</td>
<td>+</td>
<td>+</td>
<td>0.00</td>
<td>0.15</td>
</tr>
<tr>
<td>9</td>
<td>+</td>
<td>+</td>
<td>0.00</td>
<td>0.10</td>
</tr>
<tr>
<td>10</td>
<td>+</td>
<td>ND</td>
<td>0.00</td>
<td>ND</td>
</tr>
<tr>
<td>11</td>
<td>+</td>
<td>+</td>
<td>0.15</td>
<td>0.10</td>
</tr>
<tr>
<td>12</td>
<td>+</td>
<td>+</td>
<td>0.00</td>
<td>0.05</td>
</tr>
<tr>
<td>13</td>
<td>+</td>
<td>+</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>14</td>
<td>+</td>
<td>+</td>
<td>0.00</td>
<td>0.30</td>
</tr>
<tr>
<td>15</td>
<td>+</td>
<td>+</td>
<td>17.30</td>
<td>1.00</td>
</tr>
<tr>
<td>16</td>
<td>+</td>
<td>+</td>
<td>0.10</td>
<td>1.15</td>
</tr>
<tr>
<td>17</td>
<td>+</td>
<td>+</td>
<td>17.00</td>
<td>0.15</td>
</tr>
<tr>
<td>18</td>
<td>+</td>
<td>+</td>
<td>3.35</td>
<td>0.45</td>
</tr>
<tr>
<td>19</td>
<td>+</td>
<td>+</td>
<td>2.00</td>
<td>0.30*</td>
</tr>
<tr>
<td>20</td>
<td>+</td>
<td>+</td>
<td>24.00</td>
<td>0.30</td>
</tr>
<tr>
<td>21</td>
<td>+</td>
<td>+</td>
<td>21.00</td>
<td>3.00</td>
</tr>
<tr>
<td>22</td>
<td>+</td>
<td>+</td>
<td>0.00</td>
<td>1.15</td>
</tr>
<tr>
<td>23</td>
<td>+</td>
<td>+</td>
<td>5.00</td>
<td>0.30</td>
</tr>
<tr>
<td>24</td>
<td>+</td>
<td>+</td>
<td>0.00</td>
<td>1.45*</td>
</tr>
<tr>
<td>25</td>
<td>+</td>
<td>+</td>
<td>0.00</td>
<td>6.00</td>
</tr>
<tr>
<td>26</td>
<td>+</td>
<td>+</td>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>27</td>
<td>+</td>
<td>+</td>
<td>16.30</td>
<td>0.15</td>
</tr>
<tr>
<td>28</td>
<td>ND</td>
<td>ND</td>
<td>0.00</td>
<td>1.50</td>
</tr>
<tr>
<td>29</td>
<td>+</td>
<td>+</td>
<td>2.00</td>
<td>0.30</td>
</tr>
<tr>
<td>30</td>
<td>+</td>
<td>+</td>
<td>7.30</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Data are derived from the medical charts. ND indicates not documented.

*No extubation on parental request.

TABLE 3  Considerations to Decide to Withhold or Withdraw Treatment Because of a PP in All Newborns in Group 2 (n = 16)

<table>
<thead>
<tr>
<th>Considerations</th>
<th>Hospital A (n = 11), Frequency (Patient)</th>
<th>Hospital B (n = 5), Frequency (Patient)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment does not contribute to medical condition</td>
<td>3 (17, 23, 28)</td>
<td>4 (19, 20, 29, 30)</td>
</tr>
<tr>
<td>Treatment is disproportionate</td>
<td>6 (15–18, 22, 28)</td>
<td>1 (30)</td>
</tr>
<tr>
<td>Predicted very low quality of life</td>
<td>6 (15, 16, 21, 22, 25, 27)</td>
<td>5 (19, 20, 25, 29, 30)</td>
</tr>
<tr>
<td>Predicted inability to communicate</td>
<td>3 (22, 24, 27)</td>
<td>2 (26, 30)</td>
</tr>
<tr>
<td>Predicted lack of self-sufficiency</td>
<td>1 (24)</td>
<td>1 (26)</td>
</tr>
<tr>
<td>Expected hospital dependence</td>
<td>1 (16)</td>
<td>1 (26)</td>
</tr>
<tr>
<td>Long life expectancy</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Predicted suffering</td>
<td>5 (15, 16, 22, 25, 27)</td>
<td>4 (18, 19, 29, 30)</td>
</tr>
<tr>
<td>From pain (now and future)</td>
<td>—</td>
<td>1 (30)</td>
</tr>
<tr>
<td>From discomfort</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>From functional disability</td>
<td>1 (25)</td>
<td>1 (29)</td>
</tr>
<tr>
<td>From poor prognosis</td>
<td>2 (15, 16)</td>
<td>5 (19, 20, 26, 29, 30)</td>
</tr>
<tr>
<td>From hopelessness</td>
<td>3 (15, 16, 22)</td>
<td>2 (19, 29)</td>
</tr>
<tr>
<td>Other patient not aware of own existence</td>
<td>1 (21)</td>
<td>—</td>
</tr>
</tbody>
</table>

Data are derived from interviews with the responsible neonatologists.

DISCUSSION

The attitude of neonatologists in the Netherlands regarding EoL issues has been reported extensively in several publications. This study is the first to report in detail to which practices this attitude really leads. We acknowledge several limitations of this study. First, our data may not be representative of the medical practice regarding sick newborns nationwide, because we evalu-
ated data from 2 of 10 hospitals with NICUs. However, these 10 NICUs have regular meetings, and they use the same practice guidelines including those regarding EoL decision-making. Second, no information is provided on other potentially important factors in the decision-making, such as the parents’ experience and their perception of the management of symptoms around the time of death. Third, our retrospective analysis of decision-making may have been influenced by inaccuracy in recall of arguments by the neonatologists who were interviewed. We found that 28 (93%) of a total of 30 deaths were attributable to the decision to withhold or withdraw life-sustaining treatment. Withdrawal of life-sustaining treatment was much more common than withholding treatment (86% vs 14%). The proportion of deaths that resulted from withholding or withdrawal of treatment in our study is substantially higher than that described in the early articles on neonatal EoL care (14%–30%). Most more recent reports from centers in the United States, the United Kingdom, Australia, and Europe have reported rates between 58% and 75%. Only 2 studies have described similar proportions as our study. Barton and Hodgman reported that 124 (86%) of 146 deaths had treatment withheld or withdrawn in their unit between 1998 and 2002. Arlettaz et al reported that in 93% of 199 deaths, treatment was withheld or withdrawn. The relatively high proportion in our study is likely to reflect the prevailing approach of the Dutch neonatologists that in sick newborns, it is not only the life-ending decision but also the life-prolonging decision that must be justified. In their opinion, if treatment is medically futile, then it should be stopped to prevent unnecessary suffering of the infant. Considering that this is a nationwide approach, it can be assumed that the high proportion reported by us is representative of the whole country. It may also reflect the philosophy of

<table>
<thead>
<tr>
<th>Patient</th>
<th>Classification</th>
<th>Medication Before the Last EoL Decision (Dosage)</th>
<th>Additional Medication After Last EoL Decision (Dosage)</th>
<th>Motivation for Additional Medication</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>3</td>
<td>NCTS</td>
<td>Fentanyl (5), lorazepam (B: 100)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>4</td>
<td>NCTS</td>
<td>Morphone (20)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>5</td>
<td>NCTS</td>
<td>Morphone (10)</td>
<td>Morphone (20)</td>
<td>Pain</td>
</tr>
<tr>
<td>6</td>
<td>NCTS</td>
<td>Morphone (20)</td>
<td>Morphone (20)</td>
<td>ND</td>
</tr>
<tr>
<td>7</td>
<td>NCTS</td>
<td>Morphone (15), midazolam (0.05)</td>
<td>Morphone (20)</td>
<td>Pain, dyspnea</td>
</tr>
<tr>
<td>8</td>
<td>NCTS</td>
<td>Morphone (15)</td>
<td>Morphone (20)</td>
<td>—</td>
</tr>
<tr>
<td>9</td>
<td>NCTS</td>
<td>Morphone (10)</td>
<td>Morphone (20)</td>
<td>—</td>
</tr>
<tr>
<td>10</td>
<td>NCTS</td>
<td>Morphone (10)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>11</td>
<td>NCTS</td>
<td>Fentanyl (3), midazolam (0.02)</td>
<td>Fentanyl (B: 4)</td>
<td>Pain, discomfort</td>
</tr>
<tr>
<td>12</td>
<td>NCTS</td>
<td>Fentanyl (3)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>13</td>
<td>NCTS</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>14</td>
<td>NCTS</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>15</td>
<td>NCTS</td>
<td>Morphone (5), midazolam (0.05)</td>
<td>Morphone (15), midazolam (B: 0.1)</td>
<td>Discomfort, dyspnea</td>
</tr>
<tr>
<td>16</td>
<td>NCTS</td>
<td>Morphone (15)</td>
<td>Morphone (20), midazolam (0.04)</td>
<td>Discomfort</td>
</tr>
<tr>
<td>17</td>
<td>NCTS</td>
<td>Morphone (15), midazolam (0.02), lorazepam (B: 0.1)</td>
<td>Morphone (20), midazolam (0.04)</td>
<td>Discomfort</td>
</tr>
<tr>
<td>18</td>
<td>NCTS</td>
<td>Morphone (20)</td>
<td>Morphone (50)</td>
<td>Discomfort, gasping</td>
</tr>
<tr>
<td>19</td>
<td>NCTS</td>
<td>Fentanyl (2)</td>
<td>Fentanyl (B: 5 + 5)</td>
<td>Pain, discomfort</td>
</tr>
<tr>
<td>20</td>
<td>NCTS</td>
<td>Fentanyl (2), midazolam (0.02)</td>
<td>Midazolam (I: 0.15)</td>
<td>Prevention of parental discomfort (gasping)</td>
</tr>
<tr>
<td>21</td>
<td>PP</td>
<td>Morphone (20), midazolam (0.02)</td>
<td>Vecuronium (B: 50)</td>
<td>Discomfort, parent request (gasping)</td>
</tr>
<tr>
<td>22</td>
<td>PP</td>
<td>Midazolam (I: 0.15)</td>
<td>Vecuronium (B: 50)</td>
<td>Discomfort, prevent gasping</td>
</tr>
<tr>
<td>23</td>
<td>PP</td>
<td>Midazolam (I: 0.15)</td>
<td>Morphone (20), vecuronium (40)</td>
<td>Discomfort, parent request (gasping)</td>
</tr>
<tr>
<td>24</td>
<td>PP</td>
<td>Morphone (10), vecuronium (40)</td>
<td>Morphone (10), midazolam (0.02), vecuronium (B: 50)</td>
<td>Discomfort</td>
</tr>
<tr>
<td>25</td>
<td>PP</td>
<td>Morphone (15), midazolam (I: 0.1)</td>
<td>Morphone (B: 10), midazolam (I: 0.1), vecuronium (B: 50)</td>
<td>discomfort, parent request (gasping)</td>
</tr>
<tr>
<td>26</td>
<td>PP</td>
<td>Fentanyl (2)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>27</td>
<td>PP</td>
<td>—</td>
<td>Midazolam (0.02), morphone (10)</td>
<td>Pain, discomfort</td>
</tr>
<tr>
<td>28</td>
<td>PP</td>
<td>Morphone (15) vecuronium (40)</td>
<td>Vecuronium (B: 50)</td>
<td>Gasping, parental request</td>
</tr>
<tr>
<td>29</td>
<td>PP</td>
<td>Fentanyl (2)</td>
<td>Vecuronium (B: 50)</td>
<td>—</td>
</tr>
<tr>
<td>30</td>
<td>PP</td>
<td>Fentanyl (3), midazolam (0.02)</td>
<td>Fentanyl (B: 3 + 3)</td>
<td>Discomfort</td>
</tr>
</tbody>
</table>

Data are as stated in the medical charts. B indicates bolus infusion; I, intermittent dose.

a At the time of the last EoL decision.

b Highest dose in 12 hours before final EoL decision, continuous intravenous administration unless indicated differently, in units as used in normal dosing. Normal dosing: morphine 25 to 50 μg/kg per hour; fentanyl 0.5 to 5.0 μg/kg per hour; midazolam 0.01 to 0.06 mg/kg per hour (intermittent 0.05 to 0.15 mg/kg per dose); lorazepam 0.1 to 0.4 mg/kg per dose; vecuronium 30 to 150 μg/kg per hour (intermittent 100 μg/kg per dose).

c Motivation in group 2 was taken from the interviews.
Dutch physicians that when a newborn is clearly dying or going to die despite treatment, all efforts must be made to let the child die in the arms of the parents, disconnected from the ventilator. In our study, the decision to withhold treatment under these circumstances was taken as an EoL decision, whereas in other studies, these cases were classified as deaths despite maximal support or classification remained unclear. This observation illustrates that comparison of the contribution of withholding or withdrawing treatment between studies is difficult because definitions vary between studies. Singh et al took withholding treatment to be equivalent to withholding cardiopulmonary resuscitation and withdrawing as withdrawing of mechanical ventilation. A much broader definition, as used by Wall and Partridge and by us, is likely to result in a higher rate of deaths after withholding or withdrawing of treatment.

Decisions that were based on quality-of-life arguments (PP) preceded death in 16 (57%) of 28 deaths in our study. Comparison of this finding with other studies is hampered by the fact that all other studies focused on the final decision that led to death. We included all EoL decisions that preceded death in our analysis and observed that in a substantial number of cases (9 [32%] of 28), >1 type of EoL decision was made. Our results show that in most cases, the first decision was the decision to withhold treatment because of the patient’s PP. Several hours to days later, it was determined that the patient had NCTS and the second decision to stop treatment was made. The newborn had shifted from PP to NCTS. One explanation for this shift could be that the first decision to withhold treatment was made to gain time without imposing additional burden, hoping that the restricted treatment would still lead to improvement. Instead of improving, the clinical situation of the child worsened over time, as a result of the underlying disease, resulting in absence of a chance of survival. Another explanation could be that the decision to withhold treatment was initially more acceptable for the parents. A third possibility could be that by withholding life-saving treatment, the physician created a situation in which survival was simply not possible. By evoking this situation, the physician prevented the more difficult discussion about quality of life. We think that our observation suggests that classification of EoL decisions in newborns can not be based on the final decision only. Evaluation and comparison of the decision-making process in NICUs and, specifically, the quality-of-life arguments at the end of life in newborns must include all previous decisions because of the possible causality between these decisions.

The neonatologists in our study used general and specific quality-of-life considerations to justify why the predicted quality of life was deemed very low. The general considerations relate to the moral principle of proportionality of treatment, suggesting that the “costs” in terms of additional burden for the patient versus the foreseeable “effects” in terms of improvement of the patient’s medical condition must be balanced. The most frequently used specific considerations were predicted suffering and the predicted inability to be engaged in any kind of verbal and nonverbal communication. These considerations seem to legitimize limitation of treatment for neonatologists in the Netherlands. Only a few authors have reported similar details regarding quality-of-life arguments. Wall and Partridge reported the prognosis for severe disabilities and the infant’s predicted suffering as the main quality-of-life concerns. Singh et al described that treatment was limited if the burden of continuing interventions outweighed the benefits of prolonging life. Recently, Hentschel et al reported PP: severe disabilities; and long-term, far-reaching therapy as considerations.

The infants in this study died relatively fast after the last decision to withhold or withdraw life-saving treatment. Death occurred after a median time of 30 and 60 minutes in NCTS and PP groups, respectively. We were interested in the role of medication with potential life-shortening effect around the time of death. We found that provision of sedation and analgesia was discussed at each decision-making meeting. The medication was administered before and after most EoL decisions to treat symptoms around the time of death (pain, dyspnea, and discomfort). Dosages have consistently remained within normal dosage levels in all cases. This finding suggests a limited role of the potentially life-shortening medication as the cause of death, although lethal adverse effects of the medication cannot be ruled out completely. Earlier studies about the use of potentially life-shortening drugs in the Netherlands focused on the intentions of the physician to differentiate between active ending of life and letting die. The problem with intentions is that they are very subjective, ambiguous, and sometimes unclear even to the physician himself or herself.

We have tried to find the medical reasons for the decision to administer medication. Neonatologists in our study consistently pointed out that the purpose was alleviation of symptoms in all cases. At the same time, they confirmed in the interviews that hastening death as a possible adverse effect of adequate palliative care would be acceptable. This is in line with other reports from neonatologists. The underlying “double effect” principle, suggesting that an action that causes a serious harm (death) can be permissible as an adverse effect of promoting some good end (relief of pain and suffering), is accepted in common medical practice of critical care and EoL care.

Pain and symptom management at the end of life is known to be of great concern to parents. A remarkable finding in this respect was that NMBs were administered shortly before death in 5 patients in the PP group. In 4 of these cases, the purpose was to prevent gasping of the
infant. In 3 of 4 cases, administration took place on explicit request of the parents. In the discussions with the medical team, the parents made it clear that they would not accept suffering and agony. The role of NMBs in symptom management at the end of life is controversial. Several studies have reported a practice in the use of these agents at the end of life in children.43–45 Most commentary on this issue has concluded that the initiation of these agents as the ventilator is being withdrawn is morally indefensible.40–42 Some have argued that the desire to comfort the patient’s family is an important consideration and that initiating neuromuscular blockade can be acceptable when the patient’s death after the withdrawal of mechanical ventilation is certain.33–45 However, others believe that the patient’s well-being is always more important than family interests. They argue that neuromuscular blockade potentially masks symptoms of pain and suffering and makes proper assessment and adequate treatment impossible.46 Our study is the second to report that parents sometimes explicitly request the use of lethal drugs for their child.47 We think that a request from parents shortly before a certain death is completely understandable because the sight of a gasping child is a potent source of stress and discomfort to all people who witness the dying newborn. Neonatologists in our study were prepared to grant the parental request. This suggests that they accepted the parents’ distress as their responsibility. It is uncertain whether the use of medication with a certain lethal effect can be legitimized by referring to this responsibility. We are convinced that robust palliative care is indicated in all EoL situations, providing parental education and support about the EoL physical signs in the dying child, supportive staff, anticipatory grief work, and bereavement service follow-up. If the parental request to administer NMBs to the dying newborn consistently persists despite all of theses measures, then we think that the requests should be granted in the presence of skilled and experienced clinicians.

The results of this study also confirm that deliberate ending of life in newborns remains a rare event, also in the Netherlands, where it is considered to be legally acceptable.48–50 In 6 months, no cases were registered in the study hospitals.

CONCLUSIONS
We report that the vast majority of deaths in 2 Dutch units were attributable to withholding or withdrawing of treatment. In most cases, the newborns had NCTS and prolonging of treatment could not be justified. In the remaining cases, the decision was based on a combination of quality-of-life considerations, mostly the predicted suffering and predicted inability of verbal and nonverbal communication. Potentially life-shortening medication played a minor role as a cause of death in the implementation. No case of deliberately ending the life of a newborn occurred.

REFERENCES
7. McElroy D. This doctor is proud to have killed four newborns. The Daily Telegraph. December 26, 2004:21


35. Partridge JC, Wall SN. Analgesia for dying infants whose life support is withdrawn or withheld. *Pediatrics.* 1997;99:76–79


Hand-Washing and Diapering Equipment Reduces Disease Among Children in Out-of-Home Child Care Centers

Jonathan B. Kotch, MD, MPH; Patricia Isbell, PhD, MPH; David J. Weber, MD, MPH; Viet Nguyen, MD, MPH; Eric Savage, MA; Elizabeth Gunn, BS; Martie Skinner, PhD; Stephen Fowlkes, MA; Jasveer Virk, MPH; Jonnell Allen, MPH

Departments of Maternal and Child Health and Epidemiology, School of Public Health; Departments of Medicine and Pediatrics, School of Medicine, and Frank Porter Graham Child Development Institute, University of North Carolina, Chapel Hill, North Carolina

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

ABSTRACT

OBJECTIVE. The objective of this study was to determine whether the installation of equipment for diaper-changing, hand-washing, and food preparation that is specifically designed to reduce the transmission of infectious agents would result in a decrease in the rate of diarrheal illness among children and their teachers in child care centers.

METHODS. Twenty-three pairs of child care centers were matched on size and star-rated license level. One member of each pair was randomly assigned to an intervention group and the other to a control group. Intervention centers received new diaper-changing, hand-washing, and food-preparation equipment, and both intervention and control centers received hygiene and sanitation training with reinforcement and follow-up as needed. Families with children in participating classrooms were called biweekly to ascertain the frequency and severity of any diarrheal illness episodes. Staff attendance was monitored, and staff hygiene and sanitation behaviors were observed and recorded monthly.

RESULTS. Although hygiene and sanitation behaviors improved in both intervention and control centers, there was a significant difference favoring the intervention centers with respect to frequency of diarrheal illness (0.90 vs 1.58 illnesses per 100 child-days in control centers) and proportion of days ill as a result of diarrhea (4.0% vs 5.0% in control centers) among the children. Staff in those same classrooms were called biweekly to ascertain the frequency and severity of any diarrheal illness episodes. Staff attendance was monitored, and staff hygiene and sanitation behaviors were observed and recorded monthly.

CONCLUSION. Diapering, hand-washing, and food-preparation equipment that is specifically designed to reduce the spread of infectious agents significantly reduced diarrheal illness among the children and absence as a result of illness among staff in out-of-home child care centers.
The number of children who are younger than 6 years in out-of-home child care in the United States has steadily increased in the past 30 years. According to the National Association for the Education of Young Children, 70% of US children are in nonparental child care and early education settings, spending at least part of their day with caregivers other than their parents and in groups of children other than their siblings.1 When children are cared for in nonfamily groups, there is an increase in the transmission of infectious agents.2 Otitis media,3–7 upper respiratory tract infection,5,8–10 and diarrhea5,10–14 are early childhood illnesses that may be acquired in this manner. These illnesses cause more morbidity15 and occur more commonly in children who are in group child care environments than in children who are reared in their own homes. Several studies have found that the incidence of illness episodes decreases with age,14,16–18 suggesting that early childhood is a “sensitive period” for contracting infectious illness. Numerous studies have also found that rates of illness in child care settings can be reduced by implementing simple hygiene measures such as a hand-washing program.19–22 Carabin et al23 demonstrated that the incidence rate of diarrhea was reduced by a hygiene training and monitoring program and that monitoring alone reduced the level of bacterial contamination on children’s and caregivers’ hands. Unlike health care providers, caregivers in child care settings are not provided extensive training and retraining in the correct method of dealing with potential pathogens.24 This deficiency is compounded by the high turnover of child care center staff and emphasizes the need for continuous training in sanitation and hygiene.

Infection-control programs that have been successful in reducing child care illnesses have had other benefits. Krilov et al24 reported that the implementation of an infection-control program resulted in downward trends in respiratory and gastrointestinal illnesses, number of physician visits, antibiotic use, and school days missed. Uhari et al25 reported similar results in a treatment group of children who had fewer infections and prescriptions for antimicrobial agents than did control subjects. Furthermore, infection-control programs in child care have reduced the costs that are incurred by parents of sick children and their employers. Cost/benefit analysis has found a net savings resulting from decreased spending on alternative child care, physician visits, medication, and costs that are associated with parents’ time lost from work.22

Although many studies have established a positive relationship between hygiene training of caregivers and the reduction of illness, few have examined the contribution of physical factors to the prevalence of diarrheal illness in child care. Deficiencies in equipment surfaces, food-preparation areas, diapering locations, and sink and toilet availability in child care settings may affect the transmission of pathogenic organisms. In fact, Laborde et al26 found that faucet handles were among the most contaminated sites in child care centers. Other surfaces that are porous, cracked, or damaged increase the likelihood that pathogens will escape disinfection and allow transmission, especially when contact with these surfaces is frequent.

The objective of this study was to determine whether the installation of diaper-changing, hand-washing, and food-preparation equipment that was specifically designed to reduce the transmission of infectious agents in child care centers would result in a decrease in the rate of diarrheal illness among children and reduce their teachers’ absences as a result of illness while controlling for caregiver hygiene training. The study was approved by the Institutional Review Board on Research Involving Human Subjects of the School of Public Health of the University of North Carolina at Chapel Hill (UNC-CH).

METHODS

The Quality Enhancement Project for Infants and Toddlers of UNC-CH was funded by the North Carolina Division of Child Development to improve the health and safety of infants and toddlers in child care facilities through its support of child care health consultation; grants to child care providers for health and safety enhancements; and provision of information, resources, and training for child care health consultants (CCHCs). A CCHC is a health professional who has interest in and experience with children, has knowledge of resources and child care regulations, and is comfortable linking child care settings with health resources and facilities that provide primarily education and social services.27

Thirteen CCHCs who were supported by the Quality Enhancement Project for Infants and Toddlers recommended 72 child care centers for participation in the study. Inclusion criteria were having an infant or toddler classroom with at least 5 infants or toddlers and a center director and staff who were willing to (1) complete all of the paperwork required by the study and (2) allow data collectors into their programs once a month. Five centers requested removal from consideration for various reasons (eg, director illness, environmental limitations). The remaining centers were matched in pairs by North Carolina’s star-rated licensing level28 and size. After matching, 23 pairs of centers located in 21 counties of North Carolina were randomly selected. From each pair, 1 center was randomly selected as the intervention center, the other as the control. All centers participated for the duration of the study. For the purpose of testing the success of randomization, 2 different statistical tests were used, depending on the nature of the variables used in the comparison. When comparing dichotomous variables such as classroom type × experimental group, we used χ2 statistics. When comparing continuously measured variables such as age × group, we used F statistics.
which are analogous to \( t \) values. No control variables are included in these descriptive comparisons.

Most of the centers had only 1 infant and 1 toddler classroom. Directors were requested to choose 1 classroom for the study. In the case of centers with >1 classroom, the center director selected for the study the infant or toddler classroom with the highest number of children of an appropriate age who would likely remain in the classroom for the entire 7 months of the study. This process resulted in 2 instances of an infant classroom in the treatment group being paired with a toddler classroom in the control group. Nevertheless, there were no statistically significant differences between the treatment and control groups with respect to infant or toddler age group \((\chi^2 = 2.30, \text{ degrees of freedom } [df] = 1.44, \ P = .13)\). Of importance, at the end of the study, the mean ages of the children in the intervention and control classrooms were similar \((21.26 \text{ and } 21.41 \text{ months, respectively})\), and the difference was not significant \((F = 0.04, \ df = 1.361, \ P = .84)\).

The diapering, hand-washing, and food-preparation equipment that was supplied for the study was unique (the Sabre Group, Inc, Winterville, NC; www.sabregroup.com/Hatteras/hatteras.collection.htm), incorporating cast polymer tabletops with impermeable, seamless surfacing for food preparation, diaper-changing, and hand-washing. In addition, automatic faucets and foot-activated, roll-out waste bins for diaper disposal minimized contact with the equipment by soiled hands, thereby reducing the potential spread of infectious agents. Providing separate equipment for food preparation, diaper-changing, and toddler hand-washing helped segregate these activities and reduce the risk for contamination. The equipment was installed in intervention centers before data collection commenced. Control centers received the same equipment at the completion of the study.

After the equipment was installed in the intervention centers, staff in all 46 centers were trained using the Keep It Clean training module. New staff were trained within 1 week of their being hired. Keep It Clean was specifically developed for the study on the basis of successful sanitation and hygiene training activities identified by the CCHCs. The training was intended to improve and standardize the hand-washing, sanitation, diapering, and food-preparation procedures in both intervention and control centers by addressing knowledge, attitudes, and behaviors of child care providers. Pretests and posttests were collected, and follow-up training was provided by each center’s CCHC whenever deficits in knowledge, attitude, or behavior were observed during monthly visits that were conducted by trained, objective data collectors.

The centers’ directors were responsible for recruiting children into the study by providing the parents or guardians of children in the selected classrooms with a written summary of the study and a verbal description of study procedures. At least 5 children at each center were recruited. Eligibility criteria included that the child be expected to remain in the classroom throughout the 7-month study period and be <36 months of age at the end of data collection and that at least 1 family contact could participate in a telephone survey in English. Siblings were allowed to participate when they also attended the study center and met the eligibility criteria. Between September 1, 2002, and January 31, 2003, cooperating center directors recruited a total of 487 potential subjects into the study. Of these, 70 lacked usable consents and an additional 11 could not be contacted. Eighteen potential subjects who were contacted had to be dropped from the study for reasons such as “left center” or “ineligible” (because of age, other, or unknown reasons). Therefore, illness and attendance data are based on at least 1 completed parent or guardian interview for 388 infants and toddlers (Fig 1).

Telephone interviewing began on December 3, 2002, while recruiting was still in progress. There was no statistically significant difference between experimental groups in the proportion of subjects recruited after interviewing began. The mean ages of children in the treatment and control groups whose parents were interviewed in the first interview cycle \((F = 1.05, \ df = 1.185, \ P = .31)\) did not differ significantly from those whose parents’ first interviews were after the first interview cycle \((F = 1.85, \ df = 1.160, \ P = .18)\). We know that, on average, control children participated in the study 125.4 days and intervention children 119.0 days. This difference was not significant \((F = 1.29, \ df = 1.369, \ P = .26)\). Neither was there any significant difference between the 2 groups in the number of subjects \((59 \text{ control and } 62 \text{ intervention})\) who were lost to follow-up \((\chi^2 = 121, \ df = 1.369, \ P = .46)\).

Children’s illnesses and child care attendance were monitored by parent or guardian telephone interview. Participating families were contacted biweekly by the Survey Research Unit of the Department of Biostatistics of UNC-CH. The family contact was asked whether, during the previous 2 weeks, the participating child(ren) (1) had attended the center; (2) had changed rooms; and (3) had experienced any illness and, if so, what the associated symptoms were. Vouchers for reduced-cost diapers that were contributed by a major supermarket chain were provided by the study to child care providers who used them to purchase diapers for use by child subjects during the course of the study, saving the parents the expense of supplying diapers for their children in child care and possibly reducing the import of pathogens from homes into both intervention and control centers.

For each of the 30 weeks of the study, a caregiver weekly attendance form was completed by the center director and mailed to the study office using a self-addressed, stamped envelope. The caregiver weekly at-
tendance form tracked the attendance of the caregivers and volunteers in the study classroom.

To ensure that sanitation and hygiene practices remained standard, field data collectors recorded baseline and 7 monthly observations of the diapering or toileting of the children and the preparation of food (including hand-washing in both cases) using a standard form, the event sampling form. This form had 8 observable caregiver behaviors for diapering/toileting and 9 behaviors for food preparation. The behaviors followed the recommended steps as presented in the *Keep it Clean* training. Most items were scored according to whether the behaviors were performed “adequately,” “inadequately,” or “not at all” on a 3-point scale (except for behaviors that logically could be scored only “yes” or “no”), and the scores were averaged. The observations were communicated to the center’s CCHC, who would visit the center, if necessary, within the subsequent 2 weeks to provide corrective guidance. The reliability of the field data collectors and the event sampling form was checked by comparing the scores of 2 data collectors who were rating the same events concurrently. Initial reliability was >85%, and reliability remained at this high level.

Before hypothesis testing, the success of random assignment of classrooms to the intervention or control conditions was assessed. The data analysis for this purpose was generated using SAS/STAT 8.02 of the SAS system for Windows. Multivariate analysis of variance was conducted to determine whether significant group differences were detected for 14 key characteristics of the centers and classrooms that could affect the outcomes: (1) teacher/child ratio, (2) center’s star-rated license type, (3) total center enrollment, (4) total classroom enrollment, (5) age of youngest child in classroom, (6) age of oldest child in classroom, (7) number of children in classroom enrolled in the study, (8) number of subsidized children in classroom, (9) number of subsidized children enrolled in study from each classroom, (10) number of boys in classroom, (11) number of boys enrolled in study from each classroom, (12) number of caregivers in classroom, (13) number of relief caregivers per week, and (14) number of potential caregivers per week. Four of the 14 variables—mean classroom enrollment (*P* < .01), mean number of children participating in the study per classroom (*P* < .05), mean number of boys enrolled in the classroom (*P* < .001), and mean number of boys participating in the study per classroom (*P* < .05)—were significantly different between intervention and control classrooms. Because the direction of the differences—more boys and more total children in intervention classrooms—would mitigate against the intervention’s succeeding, these variables did not need to be controlled for in the models (Table 1).

Incidence density scores were computed for all episodes of diarrhea (defined as any loose, watery stool that if contained would assume the shape of the container). A separate episode of diarrhea was defined by an interval of 7 diarrhea-free days. Review of the distribution of incidence density scores for all incidences of diarrhea (mild, moderate, and severe) indicated extreme skewness with an inflated proportion of no (0) incidences. Proportions for 3 additional variables—number of days child sick, number of full days child absent from child care because of illness, and number of full days parent missed work because of child’s illness—computed by dividing the number of days by the number of biweekly telephone interviews times 14 were also highly skewed with an inflated proportion of no incidences. Therefore,
to assess significant group differences in diarrheal frequency, days ill, full days absent, and full days missed from work, a Poisson regression procedure in the LIMDEP (limited dependent variable model) software package was used.31

Three characteristics of the Poisson distribution that make it appropriate for this analysis are that (1) there are no negative values, (2) the data are highly skewed, and (3) the variance increases as the mean increases. Although our outcomes were entered as raw counts, Poisson regression automatically uses a log transformation that adjusts for skewness and prevents the model from producing negative predicted values. This procedure also allows controlling for group differences in length of exposure to the intervention between children in treatment centers and children in control centers. The distribution of differences that was created by some children’s leaving during the study and others’ beginning their center attendance after the intervention had begun was controlled for by including a variable that was equal to the log transformation of the number of days of data collection for each child as a predictor in the regression models. Data from children within the same classroom were assumed to be nonindependent. An adjustment in the covariance structure was made to account for this nonindependence by estimating a random effect for classroom.

**RESULTS**

Four Poisson regression models were estimated. In each case, the predictors included the estimated intercept of the line (where all of the predictors were equal to 0), the log transformation of the number of data collection days, and the dichotomous variable for treatment versus control group. The dependent variables were (1) frequency of severe diarrhea, (2) number of days ill with diarrhea, (3) number of full days the child was absent from child care because of diarrhea, and (4) number of full days a parent missed work because of child’s illness. Maximum likelihood estimates of the effects of these predictors indicated that the children in the intervention group experienced significantly fewer episodes of diarrhea (0.90 vs 1.58 diarrhea illnesses per 100 child-days; \( P < .001 \)) and were sick with diarrhea a lower proportion of days (4.0% vs 5.0%; \( P < .001 \)) than the children in the control group. No significant differences were found between the intervention and control groups for number of full days absent from child care or number of full days parents missed work because of child’s illness.

A similar analysis using data from the caregiver weekly attendance form was conducted to determine whether caregivers who were working in the intervention sites experienced fewer sick days than those who were working in the control sites. The predictors in the model were the number of days the caregiver worked at the site, the number of days the site was open for children to attend, and the dichotomous group variable. Estimates were generated controlling for clustering by estimating a random effect for centers. In this analysis, the caregivers in intervention sites reported a significantly lower proportion of days absent from work as a result of any illness than did the caregivers in the control sites (0.77% vs 1.73%; \( P < .001 \); Table 2).

The final analysis was conducted on the event sampling data to determine whether the diapering and food-preparation behaviors of the caregivers differed in the 2 groups of classrooms during the intervention period. A score was developed from the event sampling measure. First, the reverse score for each item was averaged across the multiple events sampled at each observation session.
Then a mean across all of the observations was computed for each item. An item analysis was conducted on the 17 items, and they were found to have adequate internal consistency (Cronbach’s $\alpha = .73$). On the basis of this evidence, the items were averaged to form an overall score for diapering and food preparation for each caregiver. Review of the distribution of this variable indicated adequate normality for an analysis of variance. No group differences were detected ($F = 0.74$, $df = 1,45$, $P = .3941$; Fig 2).

### DISCUSSION

We believe this to be the first study to investigate the impact that physical equipment in child care centers may have on the occurrence and the duration of infectious illness among both children and staff. In preparation for the study, the study staff examined commercially available diaper-changing and hand-washing equipment that was actually available for use in child care centers in North Carolina. This informal investigation revealed a lack of durable, high-quality options. Additional explo-
ration revealed that only 1 manufacturer was interested in manufacturing the necessary numbers of diaper-changing tables and food-preparation surfaces that matched the quality and the durability criteria that the study team and the state child care health consultant had developed. Other manufacturers that were contacted by the study staff declined, citing a business priority for less expensive equipment that would need to be replaced every few years.

This study has shown that high-quality equipment, characterized by seamless, impermeable countertops and touchless faucets and cabinet doors, is associated with significantly fewer episodes of diarrhea among children and fewer sick days among staff. Behavioral change strategies for reducing diarrhea in out-of-home child care may be more effective if this source of contamination is controlled. Both improved staff hygiene and sanitation behavior and state-of-the-art diapering and food-preparation equipment are necessary for optimal prevention of diarrheal illness.

Two significant differences between the 2 study groups were noted. The total number of children and the number of boys were larger in the intervention classrooms. These differences may have reduced the overall effect of the intervention, because number of children per classroom is a risk factor, and boys tend to stay in diapers longer. In addition, control centers were working hard to get their perceived reward (the free equipment that they were promised at the end of the study). These 3 factors should have reduced the difference in outcomes between the intervention and control groups, suggesting that the significant differences in illnesses and absences that were found favoring the intervention group are all the more impressive.

Long-term follow-up with reinforcement of correct sanitation and hygiene behaviors resulted in steady improvement in the correct sequence of the behaviors over 7 months in both the intervention and control centers. The impact of the equipment can add value to the impact of training in proper diaper-changing and hand-washing that was observed in previous studies. Finally, an often overlooked aspect of many investigations into sanitation and hygiene in child care is the impact that infectious illness has on the teacher-caregivers and the resulting impact on the children. Ill caregivers can increase the risk to children, not just because they are vectors of disease but also because their absence results in hiring less experienced and less well-trained substitutes.

ACKNOWLEDGMENTS

This study was funded by a contract from the Division of Child Development of the North Carolina Department of Health and Human Services (June Locklear, Regulatory Services Section Chief) and supported in part by National Institute of Environmental Health Sciences grant P30ES10126.

We appreciate the contributions of Kathy Dail, Kacey Hanson, Steve Hege, Weejy Neebe, Abby Pinnix, and Billy Walton. In addition, we thank the Sabre Companies, LLC, the 13 CCHCs who work for the Quality Enhancement Project for Infants and Toddlers, and the 46 child care center directors and their staffs for their cooperation.

REFERENCES


Enzyme Replacement Therapy in Patients Who Have Mucopolysaccharidosis I and Are Younger Than 5 Years: Results of a Multinational Study of Recombinant Human α-L-Iduronidase (Laronidase)

J. Edmond Wraith, MBChB, Michael Beck, MD, Roderick Lane, PhD, Ans van der Ploeg, MD, PhD, Elsa Shapiro, PhD, ABPP, Yong Xue, MD, PhD, Emil D. Kakkis, MD, PhD, Nathalie Guffon, MD

ABSTRACT

OBJECTIVE. Our objective was to evaluate the safety, pharmacokinetics, and efficacy of laronidase in young, severely affected children with mucopolysaccharidosis I.

METHODS. This was a prospective, open-label, multinational study of 20 patients who had mucopolysaccharidosis I and were <5 years old (16 with Hurler syndrome, 4 with Hurler-Scheie syndrome) and were scheduled to receive intravenous laronidase at 100 U/kg (0.58 mg/kg) weekly for 52 weeks. Four patients underwent dosage increases to 200 U/kg for the last 26 weeks because of elevated urinary glycosaminoglycan levels at week 22.

RESULTS. Laronidase was well tolerated at both dosages. Investigators reported improved clinical status in 94% of patients at week 52. The mean urinary glycosaminoglycan level declined by ~50% at week 13 and was sustained thereafter. A more robust decrease in urinary glycosaminoglycan was observed in patients with low antibody levels and those who were receiving the 200 U/kg dosage. On examination, the liver edge was reduced by 69.5% in patients with a palpable liver at baseline and week 52 (n = 10). The proportion of patients with left ventricular hypertrophy decreased from 53% to 17%. Global assessment of sleep studies showed improvement or stabilization in 67% of patients, and the apnea/hypopnea index decreased by 5.8 events per hour (~8.5%) in those with abnormal baseline values. The younger patients with Hurler syndrome (<2.5 years) and all 4 patients with Hurler-Scheie syndrome showed normal mental development trajectories during the 1-year treatment period.

Key Words: enzyme replacement therapy, Hurler syndrome, laronidase, MPS I

Abbreviations:
MPS I—mucopolysaccharidosis I
HSCT—hematopoietic stem cell transplantation
ERT—enzyme replacement therapy
AE—adverse event
IgG—immunoglobulin G
AHI—apnea/hypopnea index
IAR—infusion-associated reaction

Accepted for publication Dec 13, 2006

Address correspondence to J. Edmond Wraith, MBChB, Royal Manchester Children’s Hospital, Wilink Biochemical Genetics Unit, Hospital Road, Pendlebury, Manchester M27 7HA, United Kingdom. E-mail: ed.wraith@cmcc.nhs.uk

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2156
doi:10.1542/peds.2006-2156
Ucopolysaccharidosis type I (MPS I) is caused by a deficiency of the lysosomal enzyme α-L-iduronidase, which leads to the progressive accumulation of the glycosaminoglycans dermatan and heparan sulfate, ultimately interfering with cell functioning and compromising virtually all tissues and organs. Historically, patients with MPS I have been classified into 3 clinical syndromes on the basis of differences in disease progression: Hurler (severe), Hurler-Scheie (intermediate), and Scheie (mild). Patients with Hurler syndrome experience cognitive decline in early childhood, whereas patients with Hurler-Scheie and Scheie syndromes have relatively mild, if any, cognitive impairment. However, the 3 phenotypes are not always clearly delineated, because there can be substantial overlap in symptom presentation and considerable variability in both the severity and the progression of the disease.

Since 1980, hematopoietic stem cell transplantation (HSCT) has been used to treat patients with Hurler syndrome, first using bone marrow and more recently using umbilical cord blood. HSCT improves many of the somatic features of MPS I and can preserve cognitive function, but bone, heart valve, and eye disease seem to be recalcitrant to treatment. The procedure requires HLA-matched donor cells, is prone to engraftment failure, and is associated with considerable morbidity and mortality, which has limited its use to severely affected patients who are early in the course of disease. Enzyme replacement therapy (ERT) with α-L-iduronidase (Aldurazyme; BioMarin Pharmaceutical, Inc, Novato, CA; and Genzyme Corp, Cambridge, MA). Laronidase was diluted in 100 mL or 250 mL of 0.9% sodium chloride injection and infused over 4 hours. Unlike in previous studies, laronidase was administered without 0.1% human serum albumin, in accordance with the European summary of product characteristics. On the basis of the results of an interim analysis that was performed on the first 13 patients, the study protocol was amended to allow the final 7 patients to receive a dosage of 200 U/kg laronidase from week 26 onward if their urinary glycosaminoglycan level at week 22 was >200 μg/mg creatinine; 4 of these patients qualified and received the higher dosage for the second half of the study. A Port-a-Cath was placed in 8 patients. To minimize possible infusion-associated reactions (IARs), all patients received an antipyretic and an antihistamine before each infusion.

**CONCLUSIONS.** Laronidase seems to be well tolerated and to provide clinical benefit in patients who have severe mucopolysaccharidosis I and are <5 years old. Enzyme replacement therapy is not curative and may not improve all affected organs and systems in individuals when irreversible changes have developed. The long-term clinical outcome and effects of antibodies and laronidase dosing on glycosaminoglycan reduction warrant additional investigation.

**METHODS**

**Study Design**

This was a 52-week, prospective, open-label, multicenter study of laronidase in 20 patients who had MPS I and were enrolled at 4 sites in the United Kingdom, France, Germany, and the Netherlands. The primary objective of the study was to evaluate the safety of laronidase in young and severely affected patients with MPS I. Secondary objectives were to evaluate the pharmacokinetics and efficacy of laronidase in this patient population.

All patients were naïve to laronidase therapy, had to be younger than 5 years at initiation of treatment, and had to have a diagnosis of MPS I confirmed by fibroblast or leukocyte α-L-iduronidase enzyme activity <10% of normal and by genotyping. Exclusion criteria included having undergone or being under consideration for HSCT, acute hydrocephalus, clinically significant organic disease not related to MPS I, administration of an investigational drug within 30 days before study enrollment, or known hypersensitivity to components of the laronidase solution. All parents or legal guardians provided written informed consent to participate in the study. The protocol was approved by each site’s independent ethics committee. The study was designed and conducted in compliance with the principles of the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use guidelines for Good Clinical Practice.

**Laronidase Treatment**

All patients initially received weekly intravenous infusions of 100 U/kg (0.58 mg/kg) laronidase (recombinant human α-L-iduronidase) from week 2 onward if their urinary glycosaminoglycan level at week 22 was >200 μg/mg creatinine; 4 of these patients qualified and received the higher dosage for the second half of the study. A Port-a-Cath was placed in 8 patients. To minimize possible infusion-associated reactions (IARs), all patients received an antipyretic and an antihistamine before each infusion.

**Evaluation of Safety**

Safety monitoring included adverse event (AE) reporting, physical examination, clinical chemistries, hematol-
ogy parameters, urinalysis, vital signs, and electrocardiograms. AEs were characterized by severity and by relationship to study drug. An infusion-associated reaction was defined as any AE that occurred from the start of an infusion to the end of the 30-minute postinfusion observation period (longer if deemed necessary by the investigator) and that was considered to be drug related by the investigator.

Antibody Testing
Antibody titers to laronidase (immunoglobulin G [IgG]) were measured every 4 weeks with the use of a modified enzyme-linked immunosorbent assay and confirmed by immunoprecipitation. IgE testing was to be performed after moderate to severe IARs.

Pharmacokinetic Assessment
Pharmacokinetic assessments were performed at weeks 1, 13, 26, and 52. Plasma laronidase activity was determined using 4-methylumbelliferyl iduronic acid as a substrate. Pharmacokinetic parameters for laronidase were calculated using standard noncompartmental methods.

Biochemical Evaluations and Efficacy Outcome Measures
Biochemical evaluations and clinical outcome measures that were relevant to this young severely affected MPS I patient population consisted of urinary glycosaminoglycan excretion, liver size, cardiac status, upper airway obstruction during sleep, growth velocity, the investigator’s global assessment, and mental development.

Urinary glycosaminoglycan excretion was determined in the first morning void using an automated dimethylmethylene blue dye-binding procedure in a central laboratory (BioMarin Pharmaceutical Inc) and expressed as micrograms per milligram of creatinine.

Liver size was evaluated by the distance of the liver edge below the right costal margin at the midclavicular line at physical examination. Quantitative volumetric measurements by MRI and/or computed tomography would have required anesthesia and were considered too high a risk for these patients.

Two-dimensional echocardiography was performed and interpreted by local cardiologists according to a centralized protocol. Left ventricular mass z scores were calculated using normative data from Children’s Hospital (Boston, MA). In addition, local cardiologists assessed the degree of left ventricular hypertrophy and valvular appearance.

Polysomnograms with resulting apnea/hypopnea index (AHI) scores (the total number of episodes of apnea and hypopnea per hour of sleep) were interpreted centrally by a single, independent sleep study expert. Upper airway obstruction during sleep should be considered as a continuum that may not fully be represented by the AHI, especially in young children. Therefore, the severity of the upper airway obstruction during sleep was also assessed as mild, moderate, or severe by the same expert using a nonlinear clinical global assessment scale based on well-defined criteria that combined clinical observation and sleep laboratory testing results.

Height and weight were measured at baseline and at weeks 13, 26, and 52 and expressed as height-for-age and weight-for-age z scores using the Centers for Disease Control and Prevention/National Center for Health Statistics clinical growth charts. z scores were also calculated for a historical control group that consisted of untreated patients with MPS I in the same age range (6–72 months) from the MPS I Registry (BioMarin and Genzyme, data on file).

Investigators provided a global assessment of the patient’s clinical status using a 7-point scale (marked, moderate, or slight decline; no change; mild, moderate, or marked improvement) before and after 13, 26, and 52 weeks of treatment with laronidase. Exploratory assessment on cognitive function was performed at baseline and weeks 26 and 52 using the Griffiths Mental Development Scales, which is validated for children from birth to 8 years and is widely used in Europe. Each patient’s mental development trajectory was calculated by plotting the patient’s mental age equivalent versus the patient’s chronological age for each study time point. The normal trajectory for mental development has the mental age corresponding to the chronological age.

Statistics
No hypothesis testing was performed in this open-label study. Continuous data were summarized by using means, medians, and ranges. Categorical data were summarized by using frequencies and distributions. All analyses were performed with the use of SAS 8.0 software (SAS Institute, Cary, NC).

RESULTS

Patients
Twenty patients who received a diagnosis of MPS I at a mean age of 1.3 years (ranging from prenatal diagnosis to 4.5 years) were enrolled in the study at a mean age of 2.9 years (range: 0.5–5.1 years; Table 1). Patients 101 to 109 were enrolled at the study site in the United Kingdom, 201 to 207 in France, 301 and 302 in Germany, and 401 to 403 at the site in the Netherlands. As assessed by the principal investigators at the study sites, 16 (80%) patients had a diagnosis of severe Hurler syndrome and 4 (20%) patients had the attenuated Hurler-Scheie syndrome. Nearly all were white (90%) and there was a slight excess of boys (60%). The mean height/length for-age z score was −1.56, and 40% of patients had short stature (less than −2 z scores). The most common mutations identified were W402X (45%) and Q70X (20%).
Of the 20 enrolled patients, 18 (90%) completed the 52-week study.

Safety

All patients experienced at least 1 AE, the majority of which were not related to laronidase but to underlying disease. The most commonly reported AEs (>50% of patients) were pyrexia, diarrhea, vomiting, cough, rhinorrhea, rhinitis, and rash. Seven (35%) patients experienced 33 IARs (any AE that occurred from the start of the infusion to the end of the 30-min postinfusion observation period and that was considered to be drug related by the investigator), the most common of which were pyrexia (12 events in 6 patients) and chills (7 events in 4 patients). Three of these were serious and occurred in patient 403 during infusion 12 (mild increase in blood pressure and heart rate, moderate decrease in oxygen saturation). The infusion was temporarily stopped, and the patient was treated with paracetamol (acetaminophen) and supplemental oxygen. The infusion was resumed 20 minutes later, and the patient experienced no additional IARs during the trial.

Of the 4 patients whose dosage was increased from 100 to 200 U/kg laronidase after week 26, 1 had no IARs at either dosage. IARs that were reported only with the 200 U/kg dosage were mild tremor (week 33) and moderate crepitations, wheezing, and respiratory distress (week 39). IARs were easily managed by reduction of the infusion rate, temporary interruption of the infusion, or administration of an antihistamine and/or antipyretic.

Two patients died during the study as a consequence of events related to their underlying disease. A 13-month-old girl with Hurler syndrome died of cardiac failure after 25 weeks of laronidase therapy, but no autopsy was performed. She had an episode of cardiac failure before the study and became cyanotic 6 weeks before her death. A 3-year-old boy with Hurler syndrome underwent surgery for bilateral hip dysplasia and died of a postsurgical complication (accidental extubation with unsuccessful reintubation and tracheostomy) at week 48.

There were no clinically meaningful changes for any of the serum chemistry, hematologic, or urinary parameters assessed, and changes in vital signs or physical examination parameters during the study were unremarkable.

All patients in this study developed IgG antibodies to laronidase with a mean time to seroconversion of 25.8 days. Antibody titers generally rose during the first 20 weeks of the study; the highest measured titer during the study was 1:204 800. After 1 year of laronidase treatment, 4 patients had titers ≥1:1600 (1 patient was seronegative), and 12 patients had titers >1:1600. There was no apparent correlation between the time to seroconversion or antibody titer and the incidence of IARs. In the 4 patients (marked with a “b” in Table 1) who received the increased dosage of laronidase, the subsequent antibody titers did not change appreciably. Two patients tested negative for IgE antibodies after moderate IARs.

### TABLE 1 Baseline Patient Characteristics

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age at Diagnosis, y</th>
<th>Age at Enrollment, y</th>
<th>Gender</th>
<th>Genotype</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>101</td>
<td>0.9</td>
<td>4.3</td>
<td>M</td>
<td>W402X/Q70X</td>
<td>Hurler</td>
</tr>
<tr>
<td>102</td>
<td>4.5</td>
<td>4.8</td>
<td>F</td>
<td>P533R/P533R</td>
<td>Hurler-Scheie</td>
</tr>
<tr>
<td>104</td>
<td>2.1</td>
<td>2.4</td>
<td>M</td>
<td>Q70X/W402X</td>
<td>Hurler</td>
</tr>
<tr>
<td>105</td>
<td>1.0</td>
<td>3.1</td>
<td>F</td>
<td>Q70X/W402X</td>
<td>Hurler</td>
</tr>
<tr>
<td>106</td>
<td>0.1</td>
<td>4.8</td>
<td>F</td>
<td>A36E/Q70X</td>
<td>Hurler-Scheie</td>
</tr>
<tr>
<td>107</td>
<td>2.3</td>
<td>2.5</td>
<td>M</td>
<td>Q236R/T338R</td>
<td>Hurler-Scheie</td>
</tr>
<tr>
<td>108</td>
<td>0.3</td>
<td>3.9</td>
<td>M</td>
<td>W402X/W402X</td>
<td>Hurler</td>
</tr>
<tr>
<td>109</td>
<td>0.6</td>
<td>3.0</td>
<td>F</td>
<td>W402X/W402X</td>
<td>Hurler</td>
</tr>
<tr>
<td>201</td>
<td>0.9</td>
<td>1.3</td>
<td>M</td>
<td>W402X/W402X</td>
<td>Hurler</td>
</tr>
<tr>
<td>202</td>
<td>0.7</td>
<td>1.3</td>
<td>M</td>
<td>W402X/c206delA</td>
<td>Hurler</td>
</tr>
<tr>
<td>203</td>
<td>2.1</td>
<td>3.3</td>
<td>M</td>
<td>W402X/Q70X</td>
<td>Hurler</td>
</tr>
<tr>
<td>204</td>
<td>2.7</td>
<td>4.0</td>
<td>F</td>
<td>W402X/P533R</td>
<td>Hurler-Scheie</td>
</tr>
<tr>
<td>205</td>
<td>0.6</td>
<td>1.3</td>
<td>M</td>
<td>c.35_4del/c.1524 + 1G→A</td>
<td>Hurler</td>
</tr>
<tr>
<td>206</td>
<td>2.1</td>
<td>2.4</td>
<td>M</td>
<td>W402X/A327P</td>
<td>Hurler</td>
</tr>
<tr>
<td>207</td>
<td>0.5</td>
<td>0.8</td>
<td>M</td>
<td>W402X/W402X</td>
<td>Hurler</td>
</tr>
<tr>
<td>301</td>
<td>0.6</td>
<td>3.6</td>
<td>M</td>
<td>W402X/W402X</td>
<td>Hurler</td>
</tr>
<tr>
<td>302</td>
<td>0.3</td>
<td>0.5</td>
<td>F</td>
<td>Y202X/Y202X</td>
<td>Hurler</td>
</tr>
<tr>
<td>401</td>
<td>1.5</td>
<td>5.1</td>
<td>F</td>
<td>Q70X/W402X</td>
<td>Hurler</td>
</tr>
<tr>
<td>402</td>
<td>2.4</td>
<td>4.7</td>
<td>M</td>
<td>A327P/L218P</td>
<td>Hurler</td>
</tr>
<tr>
<td>403</td>
<td>1.0</td>
<td>1.2</td>
<td>F</td>
<td>Q70X/Q70X</td>
<td>Hurler</td>
</tr>
<tr>
<td>Mean</td>
<td>1.3</td>
<td>2.9</td>
<td>12M/8F</td>
<td>4 Hurler-Scheie/16 Hurler</td>
<td></td>
</tr>
</tbody>
</table>

For all patient numbers, the first digit of the number is site specific. M indicates male, F, female.

a As assessed clinically by individual investigators.

b Patients who received 200 U/kg after week 26.

c Patient 401 was 5 years 1 month of age but was enrolled because of severe disease status.
Pharmacokinetics

The range of individual patient values, although variable, was reasonably consistent across 52 weeks of treatment. The mean area under the curve increased from 0.58 ± 0.59 hours/U per mL at the first infusion to 0.94 ± 0.97 hours/U per mL at week 52, but the change was not significant. There was no apparent relationship between area under the curve and antibody titer over the 52 week period. The mean half-life ranged from 0.55 hours to 1.55 hours with large variability in individual values. The mean volume of distribution corrected for body weight decreased from 0.753 ± 0.497 L/kg to 0.246 ± 0.210 L/kg during treatment. There was a trend toward a lower volume of distribution as the antibody level increased, suggesting that the binding of antibodies to laronidase keeps more laronidase in the plasma, away from the renal epithelial cells that are believed to be the source of urinary glycosaminoglycans, resulting in a smaller distribution volume. Pharmacokinetic analyses were hampered by the relatively small sample size and large intragroup variations that did not allow for any meaningful correlation analyses.

Urinary Glycosaminoglycans and Efficacy Outcome Measures

After initiation of treatment with laronidase, urinary glycosaminoglycan levels showed a sharp decline within the first 13 weeks followed by a plateau. The mean reduction in urinary glycosaminoglycan level after 52 weeks of treatment was 61.3% for all patients, 59.1% for patients who were treated with 100 U/kg throughout the study, and 67.7% for the 4 patients who received 200 U/kg after week 26 (Table 2).

At week 26, the mean percentage decrease was comparable between the 2 dosage groups. Whereas patients who were treated with 100 U/kg showed no additional change in the mean urinary glycosaminoglycan level between weeks 26 and 52, the level was further reduced by a mean of 8.6% in the 4 patients who were treated with 200 U/kg. Urinary glycosaminoglycan levels over time in individual study patients are presented in Fig 1. Urinary glycosaminoglycan levels from healthy children who were of the same age and assessed in a separate study are provided for comparison purposes.

All 20 patients had a palpable liver edge at baseline; however, at week 52, the liver edge was palpable in only 10 patients. The mean distance of the liver edge below the right costal margin in the 10 patients with a palpable liver during the entire study decreased from 6.0 cm to 1.7 cm (mean decrease: 69.5%; Table 2). Liver volumes were also assessed by the investigators as being normal or abnormal (ie, hepatomegaly). At the start of study, the liver size was abnormal in all patients, whereas after 52 weeks of treatment, the liver size had normalized in 50% (9 of 18) of assessable patients.

The number of patients who were categorized as having mild left ventricular hypertrophy by echocardiography decreased from 52.6% (10 of 19) of assessable patients at the start of the study to 16.7% (3 of 18) of assessable patients at the final study visit. At the start of study, the mean left ventricular mass was abnormally high as demonstrated by a mean z score of 3.8 (range: 0.6–8.4). At the end of study, the mean left ventricular mass z score had decreased by 0.9 (−11.3%) for the 17 patients with available data (Table 2). For the 14 patients with left ventricular mass z scores >2 at baseline, the mean reduction in z score was 1.3 (−28.7%), demonstrating that the largest decreases in left ventricular mass were observed in the patients with hypertrophy at baseline.

Six patients had normal AHI values at baseline (AHI <10 events per hour); in 4 of these patients, the AHI remained normal for the rest of the study, and in 2 patients, the AHI increased to 11.6 and 14.7 events per hour. AHI changes were assessed in the subgroup of

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Key Efficacy Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100 U/kg (n = 12)</td>
</tr>
<tr>
<td>Urinary glycosaminoglycans level, µg/mg creatinine</td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>526.5 ± 175.7</td>
</tr>
<tr>
<td>Week 26</td>
<td>224.8 ± 109.0</td>
</tr>
<tr>
<td>Week 52</td>
<td>233.1 ± 163.3</td>
</tr>
<tr>
<td>Change, baseline to week 26, %</td>
<td>−59.7 ± 13.0</td>
</tr>
<tr>
<td>Change, baseline to week 52, %</td>
<td>−59.1 ± 21.4</td>
</tr>
<tr>
<td>Liver edge BRCM (n = 10), cm*</td>
<td>6.0 ± 2.13</td>
</tr>
<tr>
<td>Baseline</td>
<td></td>
</tr>
<tr>
<td>Week 52</td>
<td></td>
</tr>
<tr>
<td>Change, baseline to week 52, %</td>
<td></td>
</tr>
<tr>
<td>Left ventricular mass z score (n = 17)</td>
<td>3.8 ± 2.23</td>
</tr>
<tr>
<td>Baseline</td>
<td></td>
</tr>
<tr>
<td>Week 52</td>
<td></td>
</tr>
<tr>
<td>Change, baseline to week 52, %</td>
<td></td>
</tr>
</tbody>
</table>

Data are shown as mean ± SD. BRCM indicates below the right costal margin.

* Calculated only for the patients who had a palpable liver during the whole study.
patients (n = 9) who had clinically significant abnormal baseline values (AHI ≥10) and available results at the end of study. In this subgroup, the mean AHI decreased from 45.3 to 39.6 events per hour, which corresponds to an 8.5% reduction. Predefined clinical significance (ie, a 25% reduction in events per hour) was reached by 5 patients (33% of the patients with available data). Of the 9 patients with abnormal baseline AHI values, 5 patients had tonsillectomies and/or adenoidectomies performed between baseline and week 11 that might have confounded the results. However, most of the AHI improvements occurred between weeks 26 and 52, suggesting that the benefit was attributable to laronidase rather than surgery.

In the 15 patients with available sleep study results at both baseline and week 52, global assessment of the sleep study data by the central expert revealed that 5 (33%) patients showed improvement (1 mild, 3 moderate, and 1 marked), 5 (33%) patients showed no change, and 5 (33%) patients showed worsening (1 marked, 1 moderate, and 3 slight).

Changes in the height and weight growth velocities of study patients were compared with cross-sectional data from an age-matched, untreated control group of patients (27 with Hurler syndrome and 17 with Hurler-Scheie syndrome) from the MPS I Registry (BioMarin and Genzyme, data on file). There was no apparent relationship between the gender of the patient and height-for-age or weight-for-age z score over the study period of 1 year. The regression line for untreated registry patients shows that patients with MPS I disease are initially taller than their age-matched normal peers during the first 2 years of life and that, as expected, the height-for-age z score declines with age (Fig 2). Seven study patients (4 with Hurler syndrome and 3 with Hurler-Scheie syndrome) showed a net increase in height-for-age z score after 52 weeks of treatment. Likewise, untreated patients showed a declining weight-for-age z score across the same age range, whereas 3 study patients (all with Hurler syndrome) showed a net in-
crease in weight-for-age z score during the 1-year study duration (data not shown).

In the exploratory evaluation, the mental development trajectories of individual study patients during the 1-year treatment period are shown in Fig 3. Development is expressed as the mental age equivalent and was measured using the Griffiths Mental Development Scales. The patients who had Hurler syndrome and were younger than 2.5 years and the 4 patients with Hurler-Scheie syndrome showed the steepest slope of development, which was similar to that of normal age-matched children. Patients who had Hurler syndrome and were older and whose development was already starting to plateau continued to show a flat developmental trajectory. For determination of whether the developmental gains that were made in this 1-year study are sustainable, especially in the young patients with Hurler syndrome, longer follow-up is needed.

The investigator’s global assessment of the patients’ clinical status revealed that 94% (17 of 18) of the patients who completed the study showed improvement (6 moderate and 11 mild) after 1 year of treatment with laronidase; the 1 remaining patient showed no change.

For determination of whether the presence of IgG antibodies to laronidase affected urinary glycosaminoglycan excretion, antibody titers at week 51 were plotted against the percentage of urinary glycosaminoglycan reduction at week 52 (Fig 4). A more robust decrease in urinary glycosaminoglycans (above the median of 63%) was observed in patients with low antibody levels (≤1:1600) and those who received the 200 U/kg dose.

**DISCUSSION**

Several clinical studies have demonstrated consistent improvement in clinical and biochemical measures of MPS I disease after treatment with laronidase, which was safe and well tolerated.4-6 This study in severely affected patients who had MPS I and were younger than 5 years shows similar results.

Many of the patients with Hurler syndrome have 2 null mutations and are not expected to make cross-reacting immunologic enzyme.1 In theory, this could lead to a stronger humoral immune response to laronidase and an increased safety risk in these patients. All patients in this study developed IgG antibodies against laronidase. This is similar to the seroconversion rate (93%) that was seen in patients who had Hurler-Scheie and Scheie syndromes and received laronidase in previous studies.4,5 Although the overall IgG levels in this study were higher and the mean time to seroconversion was shorter than in previous studies, these findings were not associated with any noticeable differences between phenotypes in terms of IARs or mean percentage of glycosaminoglycan reduction. The type and the frequency of IARs were similar to those that were seen in older patients who were treated with laronidase in the Phase 3 double-blind study.3

In this study, patients with low or absent antibody
levels had a more robust urinary glycosaminoglycan reduction than patients who had higher antibody levels and showed more variable urinary glycosaminoglycan reductions. The use of a higher dosage (200 U/kg) might overcome the effect of high antibody levels on the urinary glycosaminoglycan excretion, as shown by the more robust urinary glycosaminoglycan reductions in the patients who received the double dosage.

The pharmacokinetics of laronidase in patients who had MPS I and were ≥5 years of age were characterized in an earlier study in which patients received 100 U/kg once a week for 26 weeks rather than 52 weeks. Taking into account the small number of patients in these studies, there is reasonable agreement in the ranges for the key pharmacokinetic parameters, suggesting no major difference between patients who are <5 and ≥5 years of age.

The degree of reduction in urinary glycosaminoglycan level and hepatomegaly indicates effective clearance of accumulated glycosaminoglycan substrates in young and severely affected patients, consistent with the treatment effect that was observed previously in older patients with more attenuated disease. By comparison, bone marrow transplantation also leads to a reduction in urinary glycosaminoglycan levels to the upper limit of normal after several months. Weekly infusion of laronidase that was started 20 months after transplantation in 1 patient with mixed chimerism has normalized glycosaminoglycan excretion (N. Guffon, MD, verbal communication, 2006). Urinary glycosaminoglycan clearance was not proportionally increased by doubling the dosage of laronidase, but an increase in dosage may be beneficial in patients with persistently high urinary glycosaminoglycan levels after an initial period of treatment at the approved 100 U/kg (0.58 mg/kg) dosage. Early dosage-ranging studies that were conducted in dogs with MPS I showed a similar result with only a modest additional improvement in tissue glycosaminoglycan clearance at a dosage of 2.0 mg/kg versus 0.5 mg/kg laronidase (BioMarin, data on file).

Exploratory clinical efficacy end points also suggest that laronidase exerts positive effects on some organ systems in this young severe MPS I patient population. Cardiorespiratory dysfunction is an important cause of morbidity and mortality across the spectrum of patients with MPS I. Left-sided valvular heart disease that is caused by mitral and/or aortic valve dysplasia and primary myocardial involvement are well documented in MPS I. The left ventricular mass was significantly increased at baseline with a z score for the group of 3.8, but the ejection fraction and other echocardiogram parameters were within the limit of normal. No indirect signs of coronary occlusion could be detected, but 1 patient died from an undefined cardiac cause 1.5 months after the start of therapy. In this study, 70% of the patients who presented with left ventricular hypertrophy showed resolution after 1 year of treatment. This is consistent with other reports on a small number of patients that have shown regression of ventricular hypertrophy in the short term (5 years) after bone marrow transplantation. The largest decreases in left ventricular mass were observed in the patients with hypertrophy at baseline.

Resolution of cardiomyopathy after laronidase treatment also has been described in a young patient who had Hurler syndrome and initially was deemed ineligible for bone marrow transplantation. Once improved, the child was able to undergo successful transplantation. During this study, only small changes in valvular structure and function were observed. Longer term follow-up is needed to determine whether laronidase is able to halt
or slow the progression of valvular disease and avoid the need for valve replacement surgery.

Exploratory mental development testing indicated that the patients with Hurler-Scheie syndrome had a normal to above-normal rate of cognitive growth during the 1-year study. Similarly, the younger (<2.5 years of age) patients with Hurler syndrome showed an increase in cognitive function at a rate similar to that of healthy children. In contrast, the older patients with Hurler syndrome did not show any significant gains or loss in cognition.

Because the intravenously infused enzyme is not expected to cross the blood-brain barrier in appreciable amounts at the administered dosage levels,21 it is possible that some of the developmental gains that were observed were indirectly related to improvement in overall health status. Long-term follow-up is needed to distinguish direct from indirect effects. The children were reported to benefit from the treatment as assessed by the investigator global assessment.

There are some limitations in the design and the methods that were used in this study. First, several factors precluded performing a randomized, controlled study, including the low incidence of MPS I, the rapidly progressive and fatal course of Hurler disease, the families’ unwillingness to receive HSCT as an alternative therapy, and ethical concerns about withholding a medication that was anticipated to receive regulatory approval during the study. Historical comparison was largely impossible because of the paucity of data from untreated patients with Hurler syndrome. The relatively short follow-up period hampers definitive conclusion on the impact of laronidase on the long-term disease progression. Comparisons from the University of Minnesota database suggest little difference early in cognitive development from HSCT (E. Shapiro, PhD, written communication, 2006). A group of 15 children who had MPS I and underwent HSCT and were the same age and had the same gender distribution had the same median developmental quotient of 65 as observed in our study group after 1 year of follow-up. Both groups were better than untreated children from the same database whose median developmental quotient was 50 at a median age of 3.3 years. Longer-term follow-up, particularly of the young patients with Hurler syndrome in this study, will clarify the cognitive development trajectory of patients who have Hurler syndrome and are treated with ERT.

The age range of the patients commanded choice of some nonquantitative explorative efficacy measures. Most, if not all, patients would not have been able to cooperate with sophisticated testing methods (eg, forced vital capacity, 6-minute walk test) as used in previous trials. In addition, the high risks that are related to anesthesia precluded the use of MRI. Because of the small sample size and variability, it was not possible to establish correlations between urinary glycosaminoglycan reductions and improvements in clinical end points in this trial.

CONCLUSIONS
Laronidase seems to be well tolerated and clinically beneficial in severely affected patients who have MPS I and are younger than 5 years. ERT is not curative and may not improve all affected organs and systems in individuals when irreversible changes have developed. Evidence of clinical benefit is provided by improvements in hepatomegaly, sleep disorder, left ventricular mass, growth velocity, mental development, and the investigator’s global assessment after 1 year of laronidase therapy. The impact on serious morbidity and mortality will require the long-term observation of a larger number of patients. Such data are best collected through the MPS I Registry, which is gathering data on all patients across the spectrum regardless of treatment status to understand better the natural history of the disease and to determine the long-term impact of ERT.

The age at diagnosis, the severity of existing symptoms, the expected disease progression, and the availability of a donor in the case of transplantation are factors to take into account when discussing treatment options with parents. In this trial, all parents and guardians were fully informed of the risk/benefit profiles of HSCT and laronidase. A variety of reasons led parents to choose ERT for their children who were younger than 5 years: delayed diagnosis (after 2 years of age), significant developmental delay, lack of a compatible HSCT donor, non-Hurler phenotype (no expected neurocognitive regression), and concern over the safety risk of transplantation. MPS I is a rapidly progressive disorder; as for other lysosomal disorders, initiation of treatment as early as possible in the disease course is required to prevent and/or minimize irreversible damage.

ACKNOWLEDGMENTS
This study was sponsored by BioMarin Pharmaceutical Inc and Genzyme Corp.

We acknowledge the participation of study patients and their families and the expert assistance of all study-site coordinators and personnel.

REFERENCES
Successful Intermittent Prophylaxis With Trimethoprim/Sulfamethoxazole 2 Days per Week for Pneumocystis carinii (jiroveci) Pneumonia in Pediatric Oncology Patients

Susan Lindemulder, MD, Edythe Albano, MD

Center for Cancer and Blood Disorders, University of Colorado Health Science Center, Children’s Hospital, Denver, Colorado

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

ABSTRACT

OBJECTIVE. This study was conducted to determine the efficacy of dosing trimethoprim/sulfamethoxazole on 2 consecutive days per week for the prevention of Pneumocystis carinii (jiroveci) pneumonia in a pediatric leukemia and lymphoma population and to determine whether trimethoprim/sulfamethoxazole contributes to neutropenia during maintenance therapy.

METHODS. Charts were reviewed for all pediatric patients with leukemia and lymphoma diagnosed between January 1, 1993, and December 31, 2002. Data were collected through April 1, 2004.

RESULTS. A total of 575 charts were reviewed; 529 patients were included in the analysis. A total of 482 (345 leukemia, 137 lymphoma) patients were evaluated on trimethoprim/sulfamethoxazole dosed 2 consecutive days per week for 268,074 patient-days. No breakthrough cases were documented in compliant patients; 2 noncompliant patients developed P carinii pneumonia. A total of 238 patients who were on trimethoprim/sulfamethoxazole prophylaxis and 13 patients who were receiving an alternative medication prophylaxis were evaluated for neutropenia during maintenance therapy. The median number of maintenance days on trimethoprim/sulfamethoxazole was 605.5 days and on alternative drug was 617 days. The median number of neutropenic maintenance days on trimethoprim/sulfamethoxazole was 15.5 days and on the alternative drug was 16 days. The median proportion of neutropenic days per patient was 0.029 on trimethoprim/sulfamethoxazole and 0.022 on the alternative drug.

CONCLUSIONS. Intermittent dosing of trimethoprim/sulfamethoxazole on 2 consecutive days per week is an effective alternative prophylactic regimen for P carinii pneumonia in pediatric patients with leukemia and lymphoma. This analysis does not support a difference in neutropenia during maintenance therapy between patients who are treated with trimethoprim/sulfamethoxazole versus an alternative drug.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-1360
doi:10.1542/peds.2006-1360

Key Words
Pneumocystis carinii (jiroveci), chemoprophylaxis, trimethoprim/sulfamethoxazole, leukemia, lymphoma, pediatrics

Abbreviations
PCP—Pneumocystis carinii pneumonia
TMP/SMX—trimethoprim/sulfamethoxazole
ALL—acute lymphoblastic leukemia

Accepted for publication Dec 12, 2006
Address correspondence to Susan Lindemulder, MD, Children’s Hospital, Center for Cancer and Blood Disorders, 1056 E 19th Ave, B115, Denver, CO 80218. E-mail: lindemulder.susan@tchden.org

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
Pneumocystis carinii (jiroveci) pneumonia (PCP) is a widely recognized opportunistic infection in immunocompromised patients, including pediatric oncology patients. Before the initiation of routine prophylaxis for PCP with trimethoprim/sulfamethoxazole (TMP/SMX), attack rates as high as 43% were reported in subsets of pediatric oncology patients. In 1977, Hughes et al published results of the first study to document successful prophylaxis with daily TMP/SMX in pediatric oncology patients. The success of daily and intermittent prophylactic dosing of TMP/SMX has subsequently been recognized by several authors. Current recommendations for TMP/SMX dosing for PCP prophylaxis in immunocompromised patients are based on either daily dosing or dosing 3 consecutive days per week.

This study reports >10 years of experience with intermittent dosing of TMP/SMX based on a regimen of 2 consecutive days per week used routinely for PCP prophylaxis in pediatric patients with leukemia and lymphoma. This dosing regimen was adopted from dosing in bone marrow transplant patients who returned to our center from outside institutions.

METHODS
Charts were reviewed for all pediatric oncology patients who received a diagnosis of leukemia or lymphoma between January 1, 1993, and December 31, 2002. Patients were identified through the tumor registry and oncology databases that are maintained at this institution. Data gathered included patient diagnosis; regimen for PCP prophylaxis; length of time on prophylaxis; alterations to the prophylaxis regimen; interruptions to prophylaxis (including duration and reason); reason for stopping prophylaxis; and, for patients with acute lymphoblastic leukemia (ALL) or lymphoblastic lymphoma who underwent a maintenance phase of therapy, the number and the duration of neutropenic episodes during maintenance therapy. Data were collected through April 1, 2004. Data that were obtained from the medical charts regarding pathologic diagnoses and complete blood counts were confirmed against computerized laboratory reports when possible.

For uniformity in data collection, the following terms were defined: neutropenia denotes an absolute neutrophil count ≤500 cells per mm3; the prophylaxis start date is either the date of oncologic diagnosis or, when the diagnosis was made at another institution, the date on which the child was first seen at the Children’s Hospital and prophylaxis was documented; and the prophylaxis end date was the off-therapy date for the treatment protocol, the end review date of April 1, 2004, the date on which the patient went to bone marrow transplant, the date on which the patient was last seen in the clinic if care was transferred, or the date of death.

Nonparametric tests were used in the comparison of neutropenia during maintenance therapy between patients who were treated with TMP/SMX and alternative drug regimens. A Fisher’s exact test was used to compare the groups with regard to the proportion of patients with no neutropenic days observed. A Wilcoxon 2-sample test and an analysis of covariance were used to compare the proportion of neutropenic days per patient between the groups. A χ2 analysis was performed to compare the total proportion of neutropenic days for the groups as a whole.

RESULTS
Patient Characteristics and Prophylaxis Regimens
Between January 1, 1993, and December 31, 2002, 575 patients received a diagnosis of leukemia or lymphoma (409 leukemia, 166 lymphoma). The distribution of diagnoses is shown in Table 1.

Forty-six patients (33 leukemia, 13 lymphoma) were excluded from additional analysis. Reasons for exclusion were inability to locate a chart in 12 (10 leukemia, 2 lymphoma), treatment with bone marrow transplant only in 11 (all leukemia), poor documentation of the prophylaxis regimen in 15 (8 leukemia, 7 lymphoma), and death before initiation of treatment/prophylaxis in 8 (4 leukemia, 4 lymphoma). After exclusions, 529 patients remained for analysis: 376 with leukemia and 153 with lymphoma.

A total of 482 patients (345 leukemia, 137 lymphoma) received TMP/SMX prophylaxis on 2 consecutive days per week. TMP/SMX was dosed 5 mg/kg per day of TMP divided into 2 doses on 2 consecutive days per week up to a maximum dosage equivalent of 1 double-strength TMP/SMX tablet twice daily. TMP/SMX dosages were adjusted for growth throughout the course of prophylaxis.

Forty-seven patients (31 leukemia, 16 lymphoma) received PCP prophylaxis with either an alternative drug or an alternative dosing regimen of TMP/SMX. Thirty-four (25 leukemia, 9 lymphoma) received pentamidine or dapsone. Alternative drugs were used when a patient had a preexisting sulfa allergy or when the adverse effect profile of TMP/SMX was of concern. The acute myeloid

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>No. (%) of Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leukemia (n = 409)</td>
<td></td>
</tr>
<tr>
<td>ALL</td>
<td>330 (80.7)</td>
</tr>
<tr>
<td>Acute myelocytic leukemia</td>
<td>59 (14.4)</td>
</tr>
<tr>
<td>Chronic myelocytic leukemia</td>
<td>8 (2.0)</td>
</tr>
<tr>
<td>Juvenile myelomonocytic leukemia</td>
<td>5 (1.2)</td>
</tr>
<tr>
<td>Other</td>
<td>7 (1.7)</td>
</tr>
<tr>
<td>Lymphoma (n = 166)</td>
<td></td>
</tr>
<tr>
<td>Hodgkin</td>
<td>69 (41.6)</td>
</tr>
<tr>
<td>Burkitt</td>
<td>36 (21.7)</td>
</tr>
<tr>
<td>Anaplastic large cell</td>
<td>16 (9.6)</td>
</tr>
<tr>
<td>Lymphoblastic</td>
<td>17 (10.2)</td>
</tr>
<tr>
<td>Other</td>
<td>28 (16.9)</td>
</tr>
</tbody>
</table>
leukemia subgroup had the largest percentage of patients receiving alternative drugs secondary to provider bias regarding concern for myelosuppression with TMP/SMX. Thirteen (6 leukemia, 7 lymphoma) received TMP/SMX as PCP prophylaxis dosed on 3 consecutive days per week. Alternative dosing of TMP/SMX was most commonly found in patients who transferred care from another institution, and dosing was not adjusted.

**Attack Rate of PCP With TMP/SMX 2 Consecutive Days per Week**

A total of 482 patients who were receiving prophylaxis with TMP/SMX 2 consecutive days per week were evaluated for a total of 268,074 patient-days (233,716 leukemia, 34,358 lymphoma). Two (0.41%) patients developed PCP during maintenance. The median number of neutropenic days was 15.5 (range: 0–171 days). Nine (69.2%) of 13 patients who were receiving alternative medications (pentamidine, dapsone) for prophylaxis. Of the 299 patients with ALL and 14 patients with lymphoblastic lymphoma, 272 underwent a maintenance phase of therapy. Twenty-one were excluded because of medication changes during maintenance phase of therapy. Patients were included in this analysis only when they remained on either TMP/SMX or an alternative medication throughout the entire course of maintenance therapy. Twenty-one were excluded because of medication changes during maintenance therapy, leaving 238 patients on TMP/SMX and 13 on alternative medications to be evaluated for neutropenia.

**Neutropenia During Maintenance Therapy**

Patients who had a diagnosis of ALL or lymphoblastic lymphoma and were on TMP/SMX and underwent a maintenance phase of therapy were evaluated for neutropenia and compared with patients who were receiving other medications (pentamidine, dapsone) for prophylaxis. Of the 299 patients with ALL and 14 patients with lymphoblastic lymphoma, 272 underwent a maintenance phase of therapy. Patients were included in this analysis only when they remained on either TMP/SMX or an alternative medication throughout the entire course of maintenance therapy. Twenty-one were excluded because of medication changes during maintenance therapy, leaving 238 patients on TMP/SMX and 13 on alternative medications to be evaluated for neutropenia.

A total of 175 (73.5%) of 238 patients who were receiving TMP/SMX had at least 1 documented episode of neutropenia during maintenance. The median number of maintenance days on TMP/SMX was 605.5 (range: 13–1,232 days), and the median number of neutropenic days was 15.5 (range: 0–171 days). Nine (69.2%) of 13 patients who were receiving an alternative medication had at least 1 documented episode of neutropenia during maintenance. The median number of maintenance days on alternative medication was 617

<table>
<thead>
<tr>
<th>Reason</th>
<th>Leukemia Group, n</th>
<th>Lymphoma Group, n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allergy</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Inability to take medication</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>Infection</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Neutropenia</td>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>Noncompliance</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Physician choice</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>75</td>
<td>14</td>
</tr>
</tbody>
</table>

**TABLE 2  Reason for Alteration to Prophylaxis Regimen**

Changes were made to the TMP/SMX prophylaxis regimen of 2 consecutive days per week at the discretion of the primary oncologist. Eighty-nine medication changes were made in 82 patients (17.0%). The reasons for alteration are listed in Table 2. Neutropenia accounted for the majority of changes in the leukemia group (41.3%). Inability to take the medication as a result of nausea, anorexia, or mucositis accounted for the majority of alterations in the lymphoma group (64.3%). Eleven (2.3%) patients were considered to have developed a new allergy to TMP/SMX with the report of rash and had their prophylaxis drug changed. Very few changes were required secondary to noncompliance. In some instances, the prophylaxis regimen was altered but the reason was not clearly documented in the chart; these were categorized “physician choice.”

Some patients had brief interruptions to their TMP/SMX prophylaxis without actual medication change. Forty-four patients (41 leukemia, 3 lymphoma) had 57 interruptions documented. The median length of interruption was 7 days. Fifty-four (94.7%) of 57 interruptions occurred secondary to neutropenia.
(range: 273–977 days), and the median number of neutropenic days was 16 (range: 0–37 days).

Patients who were on TMP/SMX experienced neutropenia on 4.11% of maintenance days compared with 2.17% on alternative drugs ($P < .001$, $\chi^2$ test). Although previous studies reported neutropenia data using neutropenic days as a proportion of total maintenance days for all patients, given the small number of patients in the alternative drug group for this study, the 2 groups were additionally compared using the proportion of neutropenic days per patient.

The distribution of neutropenic days per patient was highly skewed as a result of the large number of patients with 0 neutropenic days (63 of 238 TMP/SMX, 5 of 13 alternative medication). This proportion of patients with 0 neutropenic days was not significantly different between the 2 groups ($P = .26$, Fisher’s exact test). The median proportion of neutropenic days was 0.029 in patients who were receiving TMP/SMX and 0.022 in those who were receiving an alternative drug. The groups were compared using a Wilcoxon 2-sample test and were not significantly different ($P = .20$). The days of neutropenia were also compared using analysis of covariance with total days on drug and drug group as covariates. This was also nonsignificant for drug group ($P = .11$).

**DISCUSSION**

*Pneumocystis* organisms were first identified in the early 20th century by Carlos Chagas while studying a model of trypanosome infection and by Antonio Carini in infected rat lung. The organism was initially thought to be a protozoan and was named *Pneumocystis carinii*. In the late 1980s, analysis of small ribosomal RNA subunits established that the organism was a fungus. The organism that causes infection in humans has been found to be different from the organism that causes infection in animals, and a change in the name to *Pneumocystis jiroveci* was proposed in the late 1990s. The name change remains controversial.

PCP was reported as an epidemic problem for children who receive treatment for malignancy in a report of 19 patients who had malignancy and were treated at St Jude Children’s Research Hospital in 1968 and 1969. In the early 1970s, PCP was described as an opportunistic infection in pediatric patients with malignancy. In 1977, Hughes established the efficacy of daily TMP/SMX in the prevention of *P carinii*-associated pneumonia in pediatric oncology patients. He followed this in 1987 by establishing the efficacy of intermittent prophylaxis with TMP/SMX given 3 consecutive days per week. PCP prophylaxis with TMP/SMX given 2 consecutive days per week has been found effective in heart transplant patients and bone marrow transplant patients.

This study establishes the efficacy of an intermittent PCP prophylaxis regimen of 2 consecutive days per week with TMP/SMX in pediatric patients with leukemia and lymphoma. Although the study is not a prospective comparison, the large number of patients and the long study period in this study support the efficacy of this practice. Our analysis documents the treatment of 482 pediatric patients with leukemia and lymphoma over ~268 000 patient-days with no breakthrough infections observed in patients who were compliant with the prophylactic regimen.

Hematologic abnormalities such as granulocytopenia and thrombocytopenia are known adverse reactions associated with use of TMP/SMX in children. These effects may be of importance in recovery after intensive chemotherapy as well as bone marrow transplant but are of particular concern in the maintenance phase of ALL therapy. In a crossover study of children with ALL in maintenance phase, patients had significant reductions in total white blood count, absolute neutrophil count, absolute lymphocyte count, and platelet count while taking daily TMP/SMX compared with placebo. Although the amount of maintenance chemotherapy in the 2 arms of this small study was similar, there remains concern that myelosuppression as a result of TMP/SMX will compromise delivery of maintenance chemotherapy.

This study compared patients who were receiving prophylaxis with TMP/SMX on 2 consecutive days per week with patients who were receiving alternative medications for prophylaxis during maintenance therapy for ALL and lymphoblastic lymphoma. Although there seemed to be a higher proportion of total days of neutropenia on TMP/SMX (4.11%) compared with alternative drug (2.17%) across all patients disregarding the exposure time for each patient, there was not a significant difference in the proportion of neutropenic days per patient. This analysis is complicated by several factors. First, the retrospective study design does not allow for uniformity in the interval at which blood counts were rechecked when found to be low. Second, although patients were evaluated only when they underwent a maintenance phase of therapy, the study period crossed multiple treatment protocols and therapy before the maintenance phase may have confounded the analysis. Finally, the size of the group that received an alternative drug compared with the group that received TMP/SMX was very small. The delivery of chemotherapy during these periods of neutropenia is likely a more important issue than the neutropenia alone. Given recommendations included in most of the treatment protocols during the study period, it is likely that patients with neutropenia as defined would have had their oral chemotherapy held; however, the retrospective nature of the study does not allow an accurate assessment of how long chemotherapy was withheld as the neutropenia resolved.

This study was not designed to address whether a
reduction to 2 days per week from 3 days per week would reduce neutropenia. However, our findings that patients had neutropenia on only 4.11% of maintenance days compares favorably with the Hughes study in which neutropenia was noted on 8.6% of maintenance days in patients who received daily prophylaxis and 5.8% of maintenance days in the 3-day intermittent prophylaxis group. Additional study is needed to compare TMP/SMX dosing on a 3 consecutive days versus 2 consecutive days versus an alternative drug regimen to explore efficacy as well as important secondary outcomes, such as the contribution to neutropenia, the impact on the delivery of chemotherapy and treatment outcomes, and the development of resistance.

ACKNOWLEDGMENT
This work was supported by the Ruth L. Kirschstein National Research Services Award for Fellowship Training in Pediatric Hematology/Oncology/Bone Marrow Transplant.

REFERENCES
Preventable and Unpreventable Causes of Childhoo-d-Onset Epilepsy Plus Mental Retardation

Carol Camfield, MD, Peter Camfield, MD

Department of Pediatrics, Dalhousie University, and the IWK Health Centre, Halifax, Nova Scotia, Canada

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

ABSTRACT

OBJECTIVE. The objective of this study was to determine the causes of childhood epilepsy associated with mental retardation and determine whether these causes are preventable.

METHODS. We selected all patients from the Nova Scotia population-based childhood epilepsy cohort (n = 692) who had mental retardation and had epilepsy onset between 1977 and 1985. Causes and family history were determined by chart review and caregiver interview after 18.8 (SD: ±7) years of follow-up.

RESULTS. Overall, 147 patients had mental retardation and epilepsy (21% of all childhood epilepsy). Standard psychological testing was available for 57%; 38.5% were too impaired for testing, which left 4% with the degree of mental retardation assessed clinically. Severe/profound mental retardation predominated (mild: 24%; moderate: 23%, severe/profound: 53%). Fifty-nine percent had additional severe neurologic deficits, most often associated with severe mental retardation. Epilepsy syndromes were symptomatic generalized (n = 73), partial (n = 58), and other (n = 16). Most had a brain imaging study: 91% had a computed tomography scan, and 12% had an MRI scan. Sixty-three percent had a defined cause: 37% had an unknown cause. A defined cause was more likely in those with severe mental retardation (60 of 78 vs 31 of 65). Identified causes were prenatal or genetic (65%), perinatal (8%), or complications of prematurity (13%). Only 11 (7%) had an acquired cause that was potentially preventable. Many (36%) had a first- or second-degree relative with epilepsy, more often in those without a clear cause (54% vs 30%) and without additional neurologic disability (57% vs 26%).

CONCLUSIONS. Approximately 20% of children with epilepsy have mental retardation. The cause is prenatal or genetic in nearly two thirds, and only 7% have an acquired, preventable cause. Important genetic influences may be present, especially in the absence of a defined cause.
Of people with mental retardation (MR), 15% to 20% have epilepsy, with the risk even higher in those with severe mental disability.\textsuperscript{1} When children with epilepsy are considered, it has been estimated that 20% to 30% have mental retardation.\textsuperscript{1–3} When epilepsy is combined with mental retardation, the chance of seizure control and remission is reduced.\textsuperscript{4,5} Even when the seizures come under control or remit, the MR means a life of dependency. In this article, we define the causes of this distressing combination and consider how often it might be potentially preventable.

**METHODS**

Patients with MR were selected from the Nova Scotia childhood epilepsy study. Details of the methods for this population-based cohort have been published elsewhere.\textsuperscript{4} In brief, the study used a central electroencephalogram reading facility as an initial case-finding method because it seems that all children with seizures in the Nova Scotia are referred for an electroencephalogram. Detailed case review with physician and hospital chart review and direct patient contact allowed identification of all of the children who developed epilepsy (2 or more unprovoked seizures) in this population between 1977 and 1985. At the time of case ascertainment, the population of Nova Scotia was ~850,000. Patients were then followed up in 2003–2005 by chart review, telephone contact with caregivers, and personal contact. For patients who could not be recontacted at that time, we used chart information up to the time of last contact. For patients who died during follow-up, chart information was used up to the time of death supplemented with telephone interviews with caregivers. Patients were excluded when they had progressive neurodegenerative metabolic disorders or brain tumors.

Mild MR was defined as a measured or estimated IQ of 50 to 70, moderate MR as IQ of 35 to 49, and severe/profound MR as IQ of <35. The degree of MR was determined by standard psychological testing or clinical judgment as late in the follow-up as possible. Epilepsy syndromes were defined as per the 1989 International League Against Epilepsy system.\textsuperscript{6} Cause was defined using the proposed system of the International League Against Epilepsy published in 2000 with some modifications. This classification system groups causes under 12 major categories. We used only 10 of these categories because we excluded patients with progressive metabolic disorders and cerebral neoplasms (Table 1). We acknowledge that some of the categories in this proposed classification are broad and not easily defined (eg, “prenatal suspected”). The cause of the MR/epilepsy combination was based on all of the information that was available up to the time of last contact. When the cause was unclear, both authors reviewed the case and arrived at a consensus. Family history of epilepsy was determined from physician charts but always confirmed directly with the parents for those whom we were able to contact in 2003–2005. The attribute “preventable” was determined by consensus and based on currently available prevention methods, rather than on only those that

### TABLE 1 Causes

<table>
<thead>
<tr>
<th>Category</th>
<th>Specific Cause</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unknown</td>
<td></td>
<td>54 (36.7)</td>
</tr>
<tr>
<td>Neurocutaneous disorders</td>
<td>Tuberous sclerosis</td>
<td>3 (5.4)</td>
</tr>
<tr>
<td>Malformation of cortical development</td>
<td>Focal heterotopia</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td>Other cerebral malformations</td>
<td>Includes neural tube defects</td>
<td>11 (7.5)</td>
</tr>
<tr>
<td></td>
<td>Aneurysm of vein of Galen</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td>Chromosomal abnormality</td>
<td>Includes trisomy 21, XXY, 15 translocation, XXX, trisomy 13, 4p-</td>
<td>6 (4.1)</td>
</tr>
<tr>
<td>Monogenic Mendelian disorders with complex</td>
<td>Includes Dravet syndrome, Rubenstein Tabi, Smith-Lemli Opitz, tricho-rhino-</td>
<td>7 (4.8)</td>
</tr>
<tr>
<td>pathogenic mechanism</td>
<td>phalangeal, arthrogryposis plus congenital heart disease</td>
<td></td>
</tr>
<tr>
<td>Prenatal or perinatal ischemic or anoxic lesions or cerebral infections causing nonprogressive encephalopathies</td>
<td>Prenatal suspected; includes congenital microcephaly, multiple nonsyndromic anomalies, small for dates (severe), congenital hydrocephalus, maternal electroshock, death of a co-twin, placenta previa, prenatal cerebral infarct, hydrencephaly</td>
<td>2 (1.4)</td>
</tr>
<tr>
<td></td>
<td>Perinatal asphyxia (severe with neonatal encephalopathy and no other cause)</td>
<td>7 (4.8)</td>
</tr>
<tr>
<td></td>
<td>Complications of prematurity</td>
<td>12 (8.2)</td>
</tr>
<tr>
<td></td>
<td>Congenital infections (TORCH [toxoplasmosis, other infections, rubella, cytomegalovirus infection, and herpes simplex])</td>
<td>2 (1.4)</td>
</tr>
<tr>
<td>Postnatal infections</td>
<td>Encephalitis</td>
<td>3 (2.0)</td>
</tr>
<tr>
<td></td>
<td>Meningitis</td>
<td>2 (1.4)</td>
</tr>
<tr>
<td>Other postnatal factors</td>
<td>Head injury</td>
<td>3 (2.0)</td>
</tr>
<tr>
<td></td>
<td>Other; includes anoxia after cardiac surgery, anoxia with bronchiolitis, intrathecal methotrexate</td>
<td>6 (4.1)</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>Primary generalized epilepsy</td>
<td>2 (1.4)</td>
</tr>
<tr>
<td></td>
<td>First-degree relative with same disorder not otherwise defined</td>
<td>2 (1.4)</td>
</tr>
</tbody>
</table>

\textsuperscript{a} The most common, but not all identified causes in each category are listed.
were available at the time of onset of epilepsy. For example, a child with epilepsy and MR from *Haemophilus influenzae* meningitis in the prevaccine era was considered to have a preventable cause.

This study was approved by the IWK Health Centre Research Ethics Board. Statistical analysis was performed by using standard $\chi^2$ and $t$ tests.

**RESULTS**

Of the 692 children who had epilepsy and make up the entire cohort, 147 (21%) were found to have MR. There were 70 girls and 77 boys with an average age of onset of epilepsy at 38 months (median: 18 months; range: 1–195 months). Forty-four percent had their first seizure in the first year of life. The average follow-up from first seizure to last contact was 18.8 ± 7 SD years. The average age at the end of follow-up was 21 ± 8 years. Twenty-nine children died during follow-up at an average age of 12 years (range: 1.4–30.0 years). The 118 survivors had an average follow-up of 20.6 ± 6 years.

The presence of MR and its degree were established by standard psychological testing in 57%. Thirty-nine percent were too impaired for testing (all severe/profound MR), leaving only 4% with the degree of MR assessed clinically by physicians. The degree of MR was significantly skewed to the more severe: 53% had severe/profound MR, 23% had moderate MR, and 24% had mild MR. In none of the patients was there clear evidence that the seizures caused the MR. Eight were judged clinically to have normal intelligence at onset and later were found to have mild MR by standardized psychological testing; however, 5 of these patients were younger than 5 years at onset, and the remaining 3 were younger than 5 years at onset. Overall, 85 (58%) had additional neurologic deficits that were judged to be severe enough to interfere with activities of daily living. Not surprising, these neurologic deficits were significantly more common in those with the most severe MR. Of those with mild, moderate, and severe/profound MR, 34%, 29%, and 81%, respectively had neurologic deficits ($P < .0001$; Table 2).

Epilepsy syndromes at onset were symptomatic localization-related (focal or partial) in 58 (39%), symptomatic (secondary) generalized in 81 (55%), and other syndromes in 8 (5%). Nearly (91%) all patients had a brain computed tomography scan, but only 12% had an MRI, reflecting the era when seizures began in this cohort.

Overall, 93 (63%) had a presumed cause, more commonly identified in those with severe/profound MR compared with lesser degrees of MR ($P < .001$). The coexistence of a severe neurologic disability also was associated with a much higher chance of identifying a cause (64% vs 18%; $P < .0001$). The cause was considered to be prenatal or genetic in 60 (64.5%), perinatal in a term infant in 7 (7.5%; all with later severe neurologic deficits), and related to complications of prematurity in 12 (12.9%). Overall, 14 (15%) were considered to have a postnatally acquired cause. Specific causes are outlined in Table 1.

A potential genetic influence for the development of epilepsy was apparent in many of these patients (Table 2). Fifty-two (35%) had a first- or second-degree relative with epilepsy, but none of these relatives had a known autosomal dominant disorder that was expressed by the proband. Patients without a defined cause for their epilepsy and MR were significantly more likely to have a positive family history for epilepsy in a first- or second-degree relative than those with a known cause (25 of 54 [46%] vs 27 of 93 [29%]; $P = .035$). This relationship strengthens when the causes of “primary generalized epilepsy” and “other genetic” are considered as cause unknown (29 of 58 [50%] vs 23 of 89 [26%]; $P = .003$).

Only 11 (7%) cases were judged to have a potentially preventable cause: 1 maternal midtrimester electrocution, 3 bacterial meningitis, 3 head injury, 1 fetal alcohol syndrome, 1 anoxia after cardiac surgery, 1 anoxia with bronchiolitis, and 1 toxicity from intrathecal methotrexate. There was no statistical relationship between the degree of MR and the presence of a preventable cause.

**DISCUSSION**

There are 3 main findings from our population-based cohort study. First, ~20% of children who develop epilepsy will have MR, most often severe/profound. Second, these children have a variety of epilepsy syndromes with a slight predominance of secondary generalized epilepsies. Third, a cause can be ascertained in nearly two thirds, usually a prenatal or genetic cause, but the cause is rarely preventable.

The rate of MR in this study is similar to that noted in the few other published epilepsy incidence studies. For example, Wakamato et al noted that 49 (32%) of 155 of long-term survivors of childhood epilepsy in 1 rural district in Japan had some degree of MR. In Finland, a 14-year follow-up of all births in 1966 in 2 provinces
showed that 21% of those with at least a single unprovoked seizure had MR. Prevalence studies tend to find higher rates of MR, reflecting that epilepsy with MR tends to be more resistant to treatment and less likely to remit.

The epilepsy and MR seemed to have the same cause, and we did not encounter children in whom the MR was clearly caused by the epilepsy. For children who were younger than 2 years, physicians have difficulty accurately assessing intelligence, particularly mild MR. Because nearly 45% of our patients had the onset of their epilepsy in the first year of life, it is difficult to decide whether the epilepsy exacerbated a preexisting intellectual deficit. For older children, especially with moderate to severe MR, we were unable to exclude a worsening of their MR from seizures because serial standardized psychometric testing was rarely available.

The finding that 36% had no clear cause for their epilepsy–MR may have been influenced by the investigations that were undertaken. Our patients developed epilepsy in an era when MRI scanning had not been developed and even computed tomography scanning was in its early days for routine clinical use. During follow-up, few patients underwent MRI scans, probably because by the time routine MRI was available, the children would have required an anesthetic for the test. In the face of stable MR, there may have been reluctance to subject the child to the risks of an elective anesthesia. The yield for unsuspected significant abnormalities on MRI in children with epilepsy and MR has not been reported; however, 3% to 4% of clinically normal children with epilepsy have been noted to have important, causative MRI findings. Nearly all of our patients were cared for at some point by a child neurologist, and the diagnosis was questioned as new information became available; however, some tests with the potential to reveal an unsuspected diagnosis, such as very complex cytogenetic analysis, were not routinely undertaken. We conclude that the rate of “no cause” could be reduced further by more extensive investigations.

The cause for the MR/epilepsy combination was prenatal or genetic in ~65% of cases. Approximately 36% had no clear cause, especially when the MR was mild. If those with no cause had been further investigated, then we suspect that any additional causes would have fallen into the prenatal or genetic categories. Genetic causes seem particularly likely because those with no clear cause had such a high rate of first- or second-degree relatives with epilepsy (~50%). Very few of these patients had a postnatal cause for their neurologic disorder, and only a tiny fraction (7%) had a prenatal or postnatal preventable cause (Table 2).

The distribution of causes for epilepsy and MR is likely in part related to socioeconomic and geographic influences. For example, tuberculous meningitis or cerebral malaria are extremely rare in Nova Scotia but common in many countries and may be a major cause of epilepsy and MR. Therefore, from the perspective of the entire world, there may be greater potential for prevention of epilepsy and MR. The perspective of Nova Scotia is that of northern North America and relative affluence. Mandatory seatbelt legislation, helmets for bicycling, high rates of immunizations, universal health care, and low levels of societal violence may have influenced the rate of preventable causes.

It is disappointing that so few cases of MR and epilepsy seem to be amenable to primary prevention. Substantial numbers of families and individuals with this difficult combination of disabilities will continue to require lifelong care.

REFERENCES
Detection and Significance of Serum Protein Marker of Hirschsprung Disease

Jia-xiang Wang, MDb, Pan Qin, MSc, Qiu-liang Liu, MSc, He-ying Yang, MSc, Ying-zhong Fan, MSc, Jie-kai Yu, PhDb, Shu Zheng, MDb

aCancer Research Institute, Zhejiang University, Hangzhou, Zhejiang, China; bPediatric Surgery Department, First Affiliated Hospital, Zhengzhou University, Zhengzhou, China

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 a specific fingerprint chromatogram model of serum proteins for early screening and diagnosis of Hirschsprung disease.

METHODS. To detect the protein mass spectrograms of 78 serum specimens (42 specimens of Hirschsprung disease, 16 specimens of adhesive ileus including appendicitis and Meckel diverticulum after operation and inflammatory bowel disease, and 20 specimens of normal control subjects), we used surface-enhanced laser desorption/ionization time of flight mass spectrometry technology, combined with bioinformatics methods (support vector machine) to develop and compare protein mass spectrograms from serum samples.

RESULTS. We identified 3 protein markers, the mass-to-charge ratio of which is positioned at 3221.7, 5639.2, and 6884.2 from the fingerprint chromatogram model of serum protein for early screening and diagnosis of Hirschsprung disease. The markers had 100% sensitivity and specificity.

CONCLUSION. The fingerprint chromatogram model of serum protein using surface-enhanced laser desorption/ionization time of flight mass spectrometry technology combining support vector machine is a new method of early screening and diagnosis of Hirschsprung disease that is worthy of additional research and application.

Key Words
Hirschsprung disease, diagnosis, SELDI, support vector machine, fingerprint chromatogram of protein

Abbreviations
HSCR—Hirschsprung disease
SELDI-TOF-MS—surface-enhanced laser desorption-ionization time-of-flight mass spectrometry
m/z—mass-to-charge ratio

Accepted for publication Dec 12, 2006
Address correspondence to Shu Zheng, MD, Cancer Institute, Zhejiang University, Hangzhou Zhejiang 310009, China. E-mail: zhengshu@mail.hz.zj.cn

www.pediatrics.org/cgi/doi/10.1542/peds.2006-1364
doi:10.1542/peds.2006-1364
Hirschsprung Disease (HSCR) is a common intestinal disease that is characterized by abnormal neurologic development of the intestine. Successful early screening and diagnosis and timely treatment significantly improve prognosis. Current clinical practice that is based on clinical suspicion alone infrequently provides early diagnosis. Advances in microprocessing and microelectronic technologies have allowed construction of microbe biochemical analysis systems on solid chip surfaces that can be used to detect quantitatively disease markers. The “protein chip” is a means of examining for the presence or absence of thousands of proteins in serum samples. The chip can be an ideal method to discover protein markers of disease.1,2 We used this technology and bioinformatic techniques to identify markers to detect whether presence or absence of specific serum proteins would distinguish children who had HSCR from normal children. Successful identification of characteristic patterns of presence or absence of specific protein peaks in this pilot study would justify future investigations using this technology in larger, prospective, population-based samples, as well as provide insights for investigations that could lead to a better understanding of the pathophysiology of HSCR.

METHODS

Materials

Seventy-eight serum specimens were obtained from the Pediatric Surgery Department of the First Affiliated Hospital of Zhengzhou University, among those, 42 case specimens were of patients with HSCR (including long-segment type [7], short-segment type [30], super short-segment type [4], and leap type [1]), 16 specimens were of adhesive ileus including appendicitis (8) and Meckel diverticulum (3) after operation and inflammatory bowel disease (5), and the 20 controls are from healthy children who had a physical checkup in Pediatric Surgery Department in our hospital. All of the specimens with HSCR were previously confirmed by pathologic results. Among the children with HSCR, 33 were male and 9 were female; their ages were 3 days to 18 years. Three cases were diagnosed when the patients presented with megacolon crisis after they were hospitalized, and they were urgently treated with colostomy. All of the blood specimens were drawn on the morning when the patients had fasted. The specimens were placed at room temperature for 1 to 2 hours and turned centrifugally for 10 minutes, and then the serum was extracted and preserved under ~80°C. Consent for collecting serum specimens was obtained from the Ethics Committee of Honan province and also from the patients and normal control subjects.

Main Reagents and Devices

CHAPS, urea, dithiothreitol, sodium acetate, and sinapinic acid were purchased from Promega (Madison, WI). PBS II+ surface-enhanced laser desorption-ionization time-of-flight mass spectrometry (SELDI-TOF-MS) and WCX2 protein chips were purchased from Ciphergen (Fremont, CA).

Protein Chip Techniques

Serum specimens were thawed in an ice bath and centrifuged at 10 000 rpm under 4°C for 2 minutes. A 96-well plate was placed on an ice box, 10 µL of U9 (9M Urea, 2% CHAPS, and 1% dithiothreitol) and 5 µL of serum were added to each well, and the plate was vibrated at 600 rpm under 4°C for 30 minutes in a cold laboratory chamber. The chip was prepared by placing it in the bioprocessor, noting the chip number, adding 200 µL of sodium acetate (100 Mm, pH 4) to each well, vibrating at 600 rpm for 2 minutes in a cold laboratory chamber, and repeating the operation once. The 96-well plate being processed by U9 was placed on the ice, and a medical gun was used to add 185 µL of sodium acetate. The plate was then vibrated at 600 rpm under 4°C for 2 minutes in a cold laboratory chamber. A total of 100 µL of processed specimens was added to the chip, which was then placed in the cold laboratory chamber under 4°C combining 600 rpm for 60 minutes. The remaining liquid was swung off, and the sample was dried rapidly. A total of 200 µL of sodium acetate was added, and after vibrating at 600 rpm for 5 minutes was swung off and dried. This was repeated 3 times. Each well was washed twice using 200 µL of deionized water, and excessive water was swung off. After the chip was air-dried, each 50% saturated sinapinic acid (1 µL) was added in 2 stages. After drying, they were placed on the device for testing.

Data Collection and Processing

A protein chip whose molecular weight was known was used to adjust the SELDI-TOF-MS system, until the tolerance of molecular weight was <0.1%. A mass spectrum reader was used to analyze the WCX2 protein chip combined with protein. Analysis parameters included laser strength of 170 and sensitivity of 6; the total number of each specimen collection was repeated 140 times. The scope of data collection was 1000 to 30 000 Da; the optimized scope was 2000 to 20 000 Da. Quality control serum was used to make repeated tests; the coefficients of variation of the peak value and the strength were 0.05% and 19.7%, respectively. All of the data used Protein Chip Software 3.1 (Ciphergen) to adjust and to make the strength and molecular weight of the total ions uniform. The ZUCI-Protein Chip Data Analyze System software package (Zhejiang University) was used to analyze the results. Peaks with mass-to-charge ratio (m/z) of <2000 were left out, and discrete wavelength analysis was used to eliminate noise and subtract the baseline. The method of local extremum was used to find out the respective peaks of the specimens and filter out the peaks with signal-to-noise ratio <2. Clustering analysis regards 10% as the minimum threshold, to cluster all of the specimens’ peaks with m/z variation <0.3% to 1 class.
Support Vector Machine
A linear support vector machine (SVM)\(^3\) classifier was used to identify peaks. Radial-based kernel function was adopted with its \(\gamma\) value set at 0.6 and penalty function \(C\) at 19. The selection of feature vector used the method of statistical filtration combined with model-dependent screening to build a discrimination model. The method of leave-1-out crossing verification was used to assess the discrimination result of the model. This experiment also used discrimination analysis methods to process mass spectrum data and to verify the result being processed by SVM.

Statistical Analysis
After the noise was filtered out of the original mass spectrum, data were filtered, and after clustering analysis, the \(m/z\) peaks data analysis was conducted using the Wilcoxon rank sum test; the testing standard was set at \(\alpha = .01\) to account for multiple testing.

RESULTS
HSCR Group and Normal Control Group
After the mass spectrum data of the HSCR group and the normal control group were filtered and screened, 213 \(m/z\) peaks were attained. After carrying out Wilcoxon rank sum tests to test relative signal strength, 13 \(m/z\) peaks with \(P < .01\) were obtained. From the random combination of protein peaks with remarkable variation, SVM screened out the combination model with the maximum Youden index of the predicted value, identifying 3 markers positioned at 3221.7, 5639.2, and 6884.2. In the HSCR group, the proteins were not significantly expressed, whereas in normal control group, they were noted to have high expression (Figs 1–3). Combining 3 potential markers, using the method of leave-1-out to make crossing detection, in the test collection of 62 patients, the specificity of discrimination model was 100%, and its sensitivity was 100%.

Adhesive Ileus Group and Normal Control Group
Through deleting 16 specimens of adhesive ileus including appendicitis and Meckel diverticulum after operation and inflammatory bowel disease and normal control group, there is no significant difference between the 2 groups positioned at the 3 markers.

DISCUSSION
HSCR is a congenital intestinal disease. The pathologic change is that a portion of the intestinal wall lacks ganglion cells. Its incidence rate is 1:5000; the proportion between male and female is 4:1. Its main cause is that ganglion cells cannot cluster and locate in the intestinal wall. Genetic predisposition and intestinal microenvironment changes may contribute to development of HSCR.\(^4\) Diagnosis of HSCR mainly relies on clinical manifestations noted by an observant caregiver and subsequent diagnostic testing using variously invasive techniques, including barium enema reduction, rectal mucus biopsy, and anorectal pressure measurement. Some newborns with HSCR are missed because their clinical signs are not typical when they are born. In newborn infants, The medical literature reports that 20% of the rectal contrasts of barium enema reduction are mistakenly diagnosed.\(^5\) This is most likely because in newborns, the dividing lines among the spasm section, transfer section, and extension section are not clear; that is, the expression of HSCR X line is not typical in this period. In addition, when carrying out barium enema reduction, because of the improper operation in the preparation of purifying intestine and injecting bubbles into the anus, “false megacolon” may be found. The barium enema is
also not without risk. If the barium is not diluted properly, then water intoxication may develop; if before examination the child has developed enterocolitis to the contrast, then enema may result in intestinal perforation and consequently lead to barium peritonitis. Another frequently used diagnostic test, rectal mucus biopsy, is a traumatic and imperfect examination. If the scope of biopsy is extensive, then the trauma is relatively large; if the scope of biopsy is narrow, then the probability of missed diagnosis increases, especially for patients with short-segment and super short-segment HSCR. Besides, transfer mucus receiving histochemical examination is at the risks of hemorrhage and perforation, and the pathologic technology requirements are relatively high. The positive rate is 83%. The diagnostic accuracy rate for anorectal pressure measurement in newborn infants with HSCR is approximately 64.3% to 71.43%; false-positives that lead to missed diagnosis may occur as a result of improper operation of the measuring device.

Susceptibility to HSCR has been attributed to the following 9 genes: RET, GDNF, NTN, EDNRB, EDN3, ECE-1, SOX-10, ZFHX1B, and PHOX2B; the ones that are more widely researched are those that are associated with abnormalities of GDNF/RET signal transmission pathways. Takahashi et al were the first to discover RET gene in the medullary thyroid cancer and discovered that it is closely related to the growth, division, transfer, and locating of intestinal ganglion cells. The research by Bordeaux et al confirmed that GDNF, which is the ligand of RET, can inhibit the cell-decaying process that is caused by RET. Warnovara et al used situ hybridization technology and reverse transcription–polymerase chain reaction method and discovered that there exist GDNF expression in the colon of both fetuses and newborn infants, but it was not discovered in other parts of intestine. Bar et al adopted an immunohistochemistry technique and discovered that GDNF was mainly present in neuroglial cells and Schwann cells, in muscular layer, and these 2 compositions are much more than that in mucosa. The theory of colonic gene expression does not explain the extremely rare sectional aganglionosis (also called “leap” HSCR), which is a rare special type. Because the clinical diagnostic tools are imperfect and a scanning of the specific candidate genes will not identify 100% of infants with HSCR, identification of ideal serum biological markers would be extremely useful. Ideally, the serum markers would identify the typical and difficult-to-diagnose nontypical, short-segment type, super short–segment type, and leap HSCR. SELDI-TOF-MS technology is a new proteome technology that was developed in 2002. It uses the principles of gene chip, reasonably combining chromatography and mass spectrum technology with protein chip, and it is able to detect proteins and peptides that are hard to verify with the traditional methods, although it does not specifically identify the proteins other than by size. It has the characteristics of rapidity, sensitivity, and high throughput, and it has made major breakthroughs in the marker screening and early diagnosis of ovarian cancer, prostate cancer, breast cancer, lung cancer, liver cancer, and colon cancer.

This experiment used SELDI-TOF-MS technology combining the method of bioinformatics. It is the first to apply proteome technology to the detection of HSCR serum protein markers and discovered specific marker proteins in the children with HSCR. These marker proteins are based on the SVM combination to build a fingerprint chromatogram model of serum protein and to discriminate successfully children with HSCR from normal children with 100% sensitivity and specificity. This diagnostic model may successfully distinguish children with HSCR from healthy children, but more testing in larger populations with well-matched case patients and control subjects and, ideally, affected and unaffected siblings is needed. It may have important practical value, particularly in the diagnosis of HSCR with nonfamilial and the nontypical types, short–segment type, super short–segment type, and the leap type. The 3 markers at the positions of 3221.7, 5639.2, and 6884.2 are of low expression in the HSCR group and of high expression in the normal control and adhesive ileus groups. This shows that we can use SELDI-TOF-MS technology to find out protein markers for the early screening and diagnosis of HSCR.

Use of SELDI-TOF-MS technology to analyze protein will produce tremendous mass spectra data that challenge the traditional data processing and analysis. SVM used in this experiment is a kind of classification technology proposed by Vapnik and others. It is a new machine learning method developed on the basis of statistical theory. In the model discrimination, the popularization, model selection, overfitting, latitude disas-
ter, and other problems of the small specimen model have been solved successfully in SVM. In the data processing of this experiment, we eliminated noise by discrete wavelength, found out mass-charge peaks of the specimens using the method of local extremum, and clustered mass-charge peaks by setting 10% as the minimum threshold. Wilcoxon rank sum test analysis assesses the relative importance of each peak in the discrimination of 2 kinds of specimens according to $P$ values. To combine randomly the remarkably different mass-charge peaks and to input them into SVM, screen out the markers, and build the discrimination model, use the method of leave 1 out to assess the model by means of crossing verification; that is, regard 1 specimen as testing and the other specimens as training, make repeated tests until the result is stable. Because each test collection is independent from the training specimens, it can be a totally blind trial. Going through these procedures and combining many methods to process the data ensures the popularization of the model built and the accuracy of the prediction.

CONCLUSIONS
The diagnostic model established in this experiment using SELDI-TOF-MS technology combined with SVM to build fingerprint chromatogram models of serum protein of patients with HSCR and normal control subjects shows its potential value in the screening and diagnosis of a difficult-to-diagnose, complex disease. Additional characterization of the identity and function of the 3 identified peptides may contribute to understanding of HSCR pathophysiology. Future studies using cases of varying phenotype and well-matched controls and prospective studies comparing the sensitivity and specificity of existing methods with this new technique are indicated.

ACKNOWLEDGMENT
We acknowledge the financial support of the National Natural Science Foundation of China (30430730).

REFERENCES
Obesity Risk for Female Victims of Childhood Sexual Abuse: A Prospective Study

Jennie G. Noll, PhDab,c, Meg H. Zeller, PhDab, Penelope K. Trickett, PhDd, Frank W. Putnam, MDe

abDepartment of Pediatrics, University of Cincinnati College of Medicine, Cincinnati Children’s Hospital Medical Center, Cincinnati, Ohio; bcDivision of Behavioral Medicine and Clinical Psychology, Center for Epidemiology and Biostatistics, and Mayerson Center for Safe and Healthy Children, University of Cincinnati, Cincinnati, Ohio; dSchool of Social Work, University of Southern California, Los Angeles, California

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

ABSTRACT

OBJECTIVE. Efforts are under way to articulate environmental, psychosocial, and biological conditions that may predispose the development and maintenance of obesity. There is increasing evidence that adverse childhood experiences such as childhood abuse may be implicated in the development of obesity. Given the dearth of prospective evidence for this link, the objective of this study was to track body mass across development (from childhood, through adolescence, and into young adulthood [ie, ages 6–27]) in a prospective, longitudinal study of abused and nonabused female subjects.

METHODS. Height and weight were obtained for 84 female subjects with substantiated childhood sexual abuse and 89 demographically similar comparison female subjects at 6 points during development. Obesity status was examined at various stages during development, and body-mass growth trajectories were contrasted across the 2 groups. It was hypothesized that, in comparison with their nonabused peers, abused female subjects would be more likely to (1) manifest obesity by early adulthood and (2) manifest high-risk growth trajectories throughout development.

RESULTS. Obesity rates were not different across groups in childhood or adolescence. By young adulthood (ages 20–27), abused female subjects were significantly more likely to be obese (42.25%) than were comparison female subjects (28.40%). Hierarchical linear modeling growth-trajectory analyses indicated that abused female subjects, on average, acquired body mass at a significantly steeper rate from childhood through young adulthood than did comparison female subjects after controlling for minority status and parity.

CONCLUSIONS. Psychosocial difficulties (eg, depression) and psychobiological conditions (eg, hypothalamic-pituitary-adrenal axis dysregulation) that have been shown to be related to both childhood abuse and obesity may help to explain these results. The identification of high-risk growth trajectories may improve health outcomes for victims. Systematic study of the mechanistic pathways and mediating processes that would help to explain the connection between childhood sexual abuse and later obesity is encouraged.
Affecting nearly 27% of adult Americans, the epidemic of obesity and its well-documented morbidity and mortality risks constitute a major public health concern. Obesity has proved difficult to prevent and treat, in part because its cause is complex and not well understood. It is estimated that 40% to 70% of the variation in obesity is heritable, leaving a substantial portion of obesity antecedents potentially characterized by relatively malleable mechanisms. Therefore, efforts are under way to articulate environmental, psychosocial, and even psychobiological conditions that may pre-dispose one to the development and maintenance of obesity. Among these efforts are increasing mandates to elucidate mechanisms that operate early in development (eg, in childhood and/or adolescence) given that obesity in this period is highly predictive of obesity in adulthood.

Research in the past 15 years has provided increasing evidence for a potential association between adverse childhood experiences and the subsequent development of obesity. In a 1974 population-based study of 756 pupils, teacher ratings of parental neglect were shown to be associated with obesity as assessed 10 years later in young adulthood. In a recent study of extremely obese adult gastric bypass candidates, 69% retrospectively reported some form of childhood maltreatment. Violations of a sexual nature are thought to be of particular salience. Sexual exploitation (eg, adulthood rape or childhood sexual molestation) was a significant correlate of adulthood obesity in a chart-review study of health maintenance organization subscribers, with 60% of victims >50 lb overweight and 25% >100 lb overweight. Adults who reported more severe childhood abuse were significantly heavier than those who reported less severe forms of abuse. Binge-eating disorder, often a comorbid condition that is seen in obese adults, has also been associated with childhood sexual abuse.

To date, there is no prospective evidence for a link between childhood sexual abuse and the subsequent development of obesity. The majority of the extant research is cross-sectional and correlational in nature, using retrospective reports of past abuse. Therefore, statements about a causal relationship between childhood abuse and obesity remain difficult to assert. This longitudinal, prospective study analyzed height and weight data over an 18-year period for 2 groups of female subjects: (1) those who were referred by child protective services (CPS) and had experienced substantiated familial sexual abuse and (2) a demographically similar group of nonabused female peers. The following hypotheses were tested: In comparison with their demographically similar nonabused peers, abused female subjects would be more likely to (1) manifest obesity by early adulthood and (2) manifest high-risk growth trajectories throughout development.

Methods

Participants

Sexually abused female subjects (N = 84) were referred by CPS agencies in the Washington, DC, metropolitan area. Eligibility criteria included (1) age 6 to 16 years, (2) participation within 6 months of disclosure, (3) substantiated sexual abuse, including genital contact and/or penetration, (4) perpetration by a family member (eg, parent, grandparent, older sibling, uncle), and (5) participation of a nonabusing caregiver (usually the biological mother). CPS records indicated that the median age at abuse onset was 7.8 years, the median duration was 24 months, 70% experienced vaginal and/or anal penetration, and 60% of perpetrators were the primary father figure (biological fathers, stepfathers, or mothers’ live-in boyfriends). These abuse characteristic were similar to comparable information reported in the 1988 National Incidence Study.

Comparison female subjects (N = 102) were recruited via advertisements in newspapers and posters in welfare, child care, and community facilities in the same neighborhoods in which the abused participants lived. Comparison families contacted study personnel and were screened for eligibility, which included having no previous contact with CPS agencies and being demographically similar to a same-aged abused female. Comparison and abused female subjects were similar in terms of residing zip codes, racial/ethnic group, age (6–16 years), predisclosure socioeconomic status (SES), family constellation (1- or 2-parent families), and other nonsexual traumatic events. At some point after entry into the study, 13 comparison female subjects revealed some form of sexual abuse and were dropped from the study, resulting in a comparison sample of 89.

Fifty-four percent of the participants were white (abused: 48; comparison: 46), 43% were black (abused: 32; comparison: 41), 2% were Hispanic (abused: 3; comparison: 2), and 1% were Asian American (abused: 0; comparison: 1). The sample ranged from low to middle SES, with mean Hollingshead scores of ~36 (defined as “blue collar,” or working class). There were no statistical differences across groups regarding mean SES or percentage of minority (ie, white versus all minority categories).

Study Design

By design, the study was cross-sequential in nature: recruiting subjects represented a cross-section of development and followed this cross-section over time longitudinally (Table 1). This design permits analyses of both static, cross-sectional within-time effects and dynamic, repeated-measures within-person effects. As illustrated in Table 1, the study began in 1987 (time 1), when participants were a mean age of 11 years. Five follow-up assessments were conducted (times 2–6). More than...
TABLE 1  Summary Statistics for the Sample and Numbers for Analyses

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Total (N = 173)</th>
<th>Abused (n = 84)</th>
<th>Comparison (n = 89)</th>
</tr>
</thead>
<tbody>
<tr>
<td>% minoritya</td>
<td>46</td>
<td>39</td>
<td>51</td>
</tr>
<tr>
<td>SES, mean ± SD (range)b</td>
<td>36 ± 12 (11–44)</td>
<td>35 ± 14 (10–47)</td>
<td>37 ± 11 (12–43)</td>
</tr>
<tr>
<td>Age at assessment (dates), mean ± SD (range)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time 3 (1990–1992)</td>
<td>13 ± 3 (8–20)</td>
<td>13 ± 3 (8–18)</td>
<td>13 ± 3 (8–20)</td>
</tr>
</tbody>
</table>

% with only 2 time points                   | 3               | 2               | 4                  |
% with only 3 time points                   | 16              | 9               | 23                 |
% with only 4 time points                   | 11              | 13              | 9                  |
% with only 5 time points                   | 22              | 34              | 12                 |
% with only 6 time points                   | 48              | 42              | 52                 |
% with only 1 time point                    | 16              | 9               | 23                 |
% with all 2 time points                    | 3               | 2               | 4                  |

96% of the sample was retained for follow-up assessments at times 4, 5, and/or 6 (abused: 82; comparison: 84). The study received approval from the university institutional review board and a federal certificate of confidentiality.

BMI and Obesity Assessments

Height and weight were obtained at each assessment by trained study personnel using a calibrated upright Health-O-Meter balance beam scale (model 400GZD; Continental Scale Corp, Bridgeview, IL). The exact apparatus and measurement procedures were used across all 6 assessments. At each assessment, participants were weighed and measured 1 time only in street clothing without shoes. These data were used to calculate BMI (kg/m²). Obesity status was defined as per Centers for Disease Control and Prevention (CDC) guidelines: CDC population BMI z scores standardized for age and gender (zBMI) ≥ 95th percentile when participants were ≤ 19 years of age and BMI ≥ 30 once participants reached age 20. Obesity status was calculated for individuals who were assessed during the following distinct developmental periods: childhood to early adolescence (ages 6–14 years), middle to late adolescence (ages 15–19 years), and young adulthood (ages 20–27 years). Female subjects received a score of 1 when they were obese at any assessment point within the age range of each developmental period and a score of 0 when they were not. BMI calculations were considered invalid when participants were pregnant at the time of an assessment. No participant fluctuated from obese to nonobese within the age range of any developmental period.

Statistical Analyses

Because of the slight sample variation in minority status and potential racial variations in BMI, minority status (1 = minority; 0 = white) was covaried in analyses. Because female subjects who have given birth were at greater risk for high BMI, parity history (0 = nulliparous; 1 = primiparous; 2 = multiparous) was an additional covariate in subsequent analyses. All analyses were performed by using SAS 9.13 (SAS Institute, Cary, NC).

With covariates included, logistic regression was used via SAS/Logistic to test for significant group main effects in obesity status and to obtain corresponding likelihood ratios of being obese given group membership (abused = 1; comparison = 0). This analysis was repeated for each of the 3 developmental periods defined, and Bonferroni α corrections were imposed accordingly (ie, α/3). Because of the cross-sequential study design, not every participant was assessed at every developmental stage. In addition, a handful of participants refused to be weighed or were pregnant at the time of assessments. Therefore, the sample size fluctuates slightly for each comparison (Table 1).
Hierarchical linear modeling (HLM)\(^1\) via SAS/Mixed was used to estimate average growth trajectories across development on the basis of raw BMI scores arrayed from ages 6 to 27. The degree to which group membership (abused = 1; comparison = 0) could account for individual variation in parameter estimates (eg, intercept, slope) was then evaluated. An advantage of HLM is that maximum likelihood estimation methods can accommodate missing data, thereby allowing the analyst to make use of all available data so that any participant with multiple time points (or ages) can be included in the analysis of the entire trajectory. Using Bayes’s estimation, individuals with more data are given more weight in the calculation, a procedure that is preferred to using listwise or pairwise deletion in analyses in which portions of the developmental curve are represented by differing individuals or any given portion of the curve is only sparingly represented.\(^1\) Restricted maximum likelihood estimation with an estimated degrees of freedom procedure\(^1\) was used to arrive at valid parameter estimates under the assumption of ignorable missing data. On the basis of population zBMI percentile scores, 22 participants (abused: 12; comparison: 10) were deemed to be obese at study entry (time 1). These 22 participants were removed from the HLM analysis to characterize growth trajectories for those who were not initially obese, thereby resulting in a total number for HLM analysis of 144. Although not necessarily contiguous, 70% of participants had at least 5 data points for the HLM growth analysis (see Table 1).**

**RESULTS**

Figure 1 includes obesity status by group comparisons in childhood/early adolescence, middle/late adolescence, and young adulthood. During childhood/early adolescence 25.42% of abused and 21.88% of comparison female subjects were obese (odds ratio [OR]: 1.25; degrees of freedom [df] = 1141; 95% confidence interval [CI]: −0.05 to 3.00; \(P = .52\)). Later in adolescence, 27.87% of abused and 15.49% of comparison female subjects were obese (OR: 2.03; \(df = 1134\); 95% CI: 0.54–4.60; \(P = .09\)). By young adulthood, 42.25% of abused and 28.40% of comparison female subjects were obese (OR: 2.85; \(df = 1151\); 95% CI: 1.06–4.64; \(P = .009\)). Results thus indicate that obesity rates were not significantly different across groups in childhood or adolescence. However, abused female subjects were 2.85 times more likely to be obese by young adulthood.

Figure 2 depicts the variation in raw BMI trajectories for all participants as well as the mean at each age from 6 to 27. These unconditional HLM results revealed a significant overall omnibus \(\chi^2 (8) = 508.79, P < .0001\) with the linear slope coefficient (1.06) significantly different from 0 (\(t_{1,244} = 10.00, P < .0001\)) and the quadratic coefficient (−0.02) significantly different from 0 (\(t_{1,245} = −4.62, P < .001\)). These results suggest that the sample, on average, showed a linear positive trend of BMI accumulation across development from childhood to young adulthood but that this trend leveled off in the early 20s.

Figure 3 shows the conditional (by group) HLM results. There was not a significant group \(\times\) intercept effect (\(t_{1,108} = 0.14, P = .89\)) or a group \(\times\) quadratic time effect (\(t_{1,108} = 0.0014, P = .79\)). These results indicate that abused and comparison female subjects did not differ with respect to BMI at intercept (age 6) or with respect...
to the quadratic trend in BMI accumulation across development (ie, the leveling off of BMI in young adulthood at approximately age 20). Results did, however, reveal a group × linear time interaction estimate (0.18) that was significantly different from 0 ($t_{1,11} = 3.03, P = .01$) with minority status and parity covaried. This result indicates that abused female subjects, on average, acquired body mass at a significantly steeper rate during development (from age 6 to age 27) than did comparison female subjects. Thus, abused female subjects, as a group, were heavier earlier in development and remained heavier throughout development than did comparison female subjects when controlling for minority status and parity. Figure 3 also depicts population statistics representing the 25th and 75th percentile trajectories as provided by the CDC (extrapolated to age 27). The average linear trend for comparison female subjects mirrors that of the population ($t_{1,244} = 0.93, P = .34$) and falls well within this boundary. However, the linear trend for abused female subjects is steeper than the CDC population trend across development ($t_{1,244} = 2.99, P = .01$) and exceeds the 75th percentile by young adulthood.

**DISCUSSION**

These results provide some of the first prospective evidence that childhood sexual abuse may place female individuals at inordinately high risk for developing and maintaining obesity. Differential obesity rates were not evident at the start of this longitudinal study, when participants were approximately the mean age of 11. However, those with histories of substantiated familial sexual abuse were more than twice as likely to be obese by young adulthood (mean age 24) as compared with their nonabused peers after controlling demographic variables and parity histories. The observed young-adulthood obesity rate in the group of sexually abused women (just over 42%) is somewhat lower than reports from other studies (eg, 69%); however, the prospective study design, stringent inclusion criteria of substantiated childhood sexual abuse, and comparison with a demographically similar nonabused peer group constitute methodologic improvements over past studies. We also observed differential rates of body-mass accrual across development, suggesting an increased likelihood that abused female individuals may manifest identifiable high-risk BMI growth trajectories from childhood through adolescence and into young adulthood.

**Caveats**

It should be made clear that we are not purporting a causal link between sexual abuse and obesity per se but are suggesting a plausible link between the various consequences that are associated with severe childhood adversity and the subsequent development of obesity. Given the documented vast individual differences in responses to childhood adversity and the potential for resilience in many victims, we are also not asserting the development of obesity as an inevitability for abuse victims. We simply wish to underscore the need for systematic study of the mechanistic and mediating processes that would help to explain the connection between childhood abuse and later obesity. Although some of our demographic matching and statistical control represent methodologic improvement over past studies, we were unable to control for several conditions that may place children at high risk for obesity development. For example, familial histories of psychiatric or substance use disorders and numerous alternative adverse environmental factors such as comorbid neglect, poor nutrition and dietary habits, family dysfunction, and greater social isolation may contribute to the development of obesity in this population independent of the experience of childhood abuse. The extent to which sexual abuse is, in itself, a particularly salient risk factor remains theoretical, because the definitive study to compare alternative forms of childhood adversity has not yet been adequately executed. Moreover, given that our sample was recruited from CPS agencies and a nonabusing caregiver was required to participate, abused participants may represent a select set of the abused population (ie, those who had reported, substantiated abuse and whose caregiver was available/supportive).

We were also unable to control for parental weight status, a known potent predictor of child and adolescent obesity that arguably is the result of a gene–environment interaction. However, our intentional omission
of initially obese individuals from the growth analysis should bolster confidence that we have presented BMI trajectories for individuals whose genetic potential for obesity may have been relatively low or unexpressed at study entry. Participant height and weight were measured only once during each assessment visit, the reliability of which could have been improved had we averaged 2 (or even 3) measurements that were obtained at each visit.

**Plausible Explanations and Future Research**

Caveats aside, these results suggest that the consequences of abuse may amplify or exacerbate the psychosocial risks that are associated with developing or maintaining obesity. For example, depression, body image disturbances, poor peer relations, low self-esteem, and the development of binge-eating disorders are psychosocial conditions that have been studied as both sequelae of abuse, and correlates of childhood and adolescent obesity. It has also been suggested that obesity might function as a defense against sexual advances, thereby serving an adaptive purpose for some victims. Research designs that include these various comorbid psychosocial conditions will be essential in any attempt to elucidate the independent contribution of childhood abuse to the development of obesity.

There may also be psychobiological conditions that develop as a result of severe childhood trauma and that may directly contribute to the development of obesity. For example, a dysregulated hypothalamic-pituitary-adrenal (HPA) axis that manifests in the hypersecretion of the stress hormone cortisol has been shown to be a response to severe childhood deprivation. Preclinical studies in rodents and primates found that early deprivation in primates is associated with permanent alterations in hormonal responses to stressors. Inordinately high catecholamine and cortisol activity has been documented in samples of abused children and adolescents. As shown in both human and animal models, HPA axis hormones play an important role in the deposition and metabolism of fat. Adrenocortical deficiency during weight loss and the high rates of obesity in Cushing syndrome underscore the role of glucocorticoids in obesity. Cortisol promotes differentiation of adipocyte precursors into adipocytes and stimulates lipogenesis in the presence of insulin. Several studies have reported a direct association between high cortisol levels and relatively large waist-to-hip ratios in women. Abnormal HPA regulation and heightened cortisol have been implicated in the development of metabolic disorders, abdominal obesity, and type 2 diabetes. Therefore, additional examination of conditions that result in HPA disruption (eg, childhood abuse) may advance our understanding of antecedents to obesity.

**CONCLUSIONS**

Although this study was not designed to give causal conclusions regarding the links between sexual abuse and obesity, the results provide a basis for the additional testing of this hypothesis. Unresolved, persistent, or untreated conditions that are commonly associated with childhood abuse (eg, depression, posttraumatic stress disorder, eating disorders) may be implicated in the development of obesity. Coping mechanisms that serve to eradicate these symptoms as well as reduce the emotional and physiologic stress associated with recovery should be encouraged. Psychological treatment of childhood abuse typically does not extend into adolescence or young adulthood, when issues that are reminiscent of the abuse (eg, the onset of dating, initiation into sexual activity) become developmentally salient. Therefore, these results suggest that childhood abuse treatment extending beyond the acute phases of recovery or that is revisited throughout development may improve health outcomes for abuse survivors.

Pediatricians who are aware of incidences of family violence may better serve patients by closely tracking BMI accumulation across development and suggesting specific treatments that might curtail obesity. With adequate child advocacy support, standard pediatric medical history intake may need to include inquiries into traumatic histories to promote optimal care. Results also suggest that female individuals with traumatic pasts may have particularly high-risk growth trajectories in late childhood and adolescence and that obesity prevention efforts targeting these points in development may be warranted. Such prevention and intervention programs that integrate both psychosocial and biological correlates of obesity may prove more efficacious than those focused on a single process.

**ACKNOWLEDGMENTS**

This research was supported by federal and foundation research grants from the National Institutes of Health (K01 HD41402 [Dr Noll] and R01 MH048330 [to Drs Trickett and Putnam]), the Department of Health and Human Services Children’s Bureau (ACYF 90CA1686 [to Drs Trickett and Noll]), the W. T. Grant Foundation, and the Smith Richardson Foundation.

We thank John L. Horn, PhD, Jeffery D. Long, PhD, Stephen R. Daniels, MD, PhD, and Brian E. Saelens, PhD, for being excellent colleagues and commenting on earlier drafts of this manuscript.

**REFERENCES**

1. Kenward MG, Roger JH. Small sample inference for fixed
3. Gunderson EP, Abrams B. Epidemiology of gestational weight
4. Hedley AA, Ogden CL, Johnson CL, Carroll MD, Curtin LR,
5. Hollingshead AF. Four Factor Index of Social Status. New Haven,
6. Williamson DF, Thompson TJ, Anda RF, Dietz WH, Felitti V.
7. Drain JG, Trickett PK, Putnam FW. Social network constellation
11. Hollingshead AF. Four Factor Index of Social Status. New Haven, CT: Department of Sociology, Yale University; 1975
33. Ng T. Studies on hormonal regulation of lipolysis and lipogenesis in fat cells of various mammalian species. Comp Biochem Physiol. 1990;97B:441–146
34. Hauner H, Schmid P, Pfeiffer EF. Glucocorticoids and insulin promote the differentiation of human adipocyte precursor cells into fat cells. J Clin Endocrinol Metab. 1987;64:832–835
Evaluating Deliberation in Pediatric Primary Care

Elizabeth D. Cox, MD, PhD, Maureen A. Smith, MD, PhD, MPH, Roger L. Brown, PhD

OBJECTIVE. Patient participation during decision-making can improve health outcomes and satisfaction, even for routine pediatric concerns. The tasks that are involved in decision-making include both information exchange and deliberation about potential options; yet deliberation (ie, the process of expressing and evaluating potential options to reach a decision) is often assessed subjectively, if at all. We objectively assessed the amount of deliberation; the involvement of parents and children in deliberation; and how deliberation is associated with child, physician, parent, and visit characteristics.

METHODS. From videotapes of 101 children’s acute care visits to 1 of 15 physicians, we coded the speaker, recipient, and timing of proposed plans (ie, options) and agreements or disagreements with the plans. Reliability of measures was assessed with Cohen’s k or intraclass correlation coefficients; validity was assessed with Spearman correlations. Outcome measures included number of plans proposed, deliberation length, and parent/child involvement in deliberation as either active (child or parent proposed a plan or disagreed with a plan) or passive (physician alone proposed plans). Multivariable models that accounted for clustering by physician were used to relate child, physician, parent, and visit factors to deliberation measures.

RESULTS. The mean number of plans proposed was 4.1, and deliberation time averaged 2.9 minutes per visit. Passive involvement of parents/children occurred in 65% of visits. After adjustment, more plans were proposed in visits by girls, and shorter deliberations occurred with college-graduate parents. Longer visits were associated with more plans proposed, longer deliberation, and reduced odds for passive parent/child involvement.

CONCLUSIONS. Using a reliable and valid technique, deliberation was demonstrated to occupy a substantial portion of the visit and include multiple proposed plans, yet passive involvement of parents and children predominated. Results support the need to develop interventions to improve parent and child participation in deliberation.
Children's participation in their health care has been demonstrated to reduce health care use, improve the quality of disease self-management, and increase child and parent satisfaction with care. A primary goal of such participation is the sharing of information to reach a treatment decision. Even for the management of an acute, commonplace condition such as otitis media, parents are more satisfied and are less likely to use antibiotics when offered the opportunity to participate in decision-making about treatment, as opposed to a less participatory approach. Furthermore, the decision-making process for many routine pediatrics issues is often influenced by parent participation and preferences, such as decisions about circumcision, vaccination for human papilloma virus, and the use of antibiotics for upper respiratory infections, yet parents and children often report feeling ill-prepared to participate in health care decisions.

Although definitions of participation in decision-making vary, the commonly cited characterization of decision-making of Charles et al clearly distinguishes between information exchange and deliberation. Information exchange includes descriptions of potential management options; fears, concerns, benefits, and risks related to these options; and patient health history, lifestyle, and social issues that are relevant to these options. Deliberation includes the expression of management options (either diagnostic or therapeutic) and preferences for or against the options. Therefore, evaluation of participation in decision-making requires assessment of both information exchange and deliberation.

Existing tools to assess participation in decision-making are limited on several fronts. First, much work has focused on assessing participation in the exchange of information during visits. Assessment of the process of expressing and evaluating potential plans (ie, the deliberative process), has been neglected, yet consideration of deliberation in addition to information exchange is critical because without deliberation, deciding on a plan of action becomes entirely a technical task. Furthermore, existing measures are often focused on physician behaviors, are subjective in nature, or are assessed by patients after the visit and therefore perhaps not reflective of actual events. Subjective measures are also suspected to be biased by dynamics of the physician–patient relationship and by both physician and patient expectations of visit participation. Four studies have objectively assessed aspects of decision-making, but assessments are limited to the occurrence of a given behavior or task and none of these studies focused on characterizing involvement in deliberation. Last, assessment of decision-making that involves a patient who is accompanied by a caregiver is rare, with no studies in pediatrics and only 1 study examining decision-making with elderly patients and their caregivers.

Wirtz et al have recognized 2 key tasks in deliberation: identifying the potential management plans from which to choose and the dialogue or negotiation about the plans. We develop objective techniques to assess these 2 deliberation tasks during pediatric primary care visits as well as the involvement of parents and children in these tasks. Such tools are important for understanding the effect of participation in deliberation on children's health outcomes for many common pediatric problems. We also apply the techniques to further our understanding of the child, physician, parent, and visit factors that influence participation in deliberation.

**METHODS**

**Participants**

To assess participation in children's primary care visits, we recruited Wisconsin-licensed family practitioners (n = 7) and pediatricians (n = 8). We selected physicians to maximize variability (~50% female, 25% minority, 50% pediatricians, 50% family physicians). After physicians were recruited, all parents who had children who were aged 3 months to 18 years and visited these physicians for an acute complaint during enrollment periods in 2001–2002 were approached before the child's medical visit. Of the 122 patients approached for participation, 101 agreed to participate (83% participation). Children who were in distress, had chronic illness, or were from non–English-speaking families were excluded.

**Data Collection**

Physicians and parents provided written consent for participation; children who were older than 7 years provided written assent. Before any patient recruitment, physicians completed a 1-page survey that included sociodemographics, specialty training and practice characteristics. Before the visit, parents completed a 1-page survey that included parent and child sociodemographics (gender; age; education; and race/ethnicity as white non-Hispanic, Hispanic, black, Asian, and other) and the child's previous health care use. Survey items were chosen on the basis of either theoretical or known associations with participation or visit outcomes. The research protocol was approved by the University of Wisconsin Human Subjects Committee and by participating sites.

Each visit was videotaped in its entirety and coded by 2 of 3 trained coders using the Noldus Observer system. Coders identified the speaker, recipient, and timing for all proposed plans and agreements/disagreements with those plans. A plan was defined as “an explicit verbal statement proposing a particular course of action for the present or the future.” Appendix 1 exemplifies the ap-
plication of our coding strategy to dialogue from 1 of our videotaped visits.

Measures
On the basis of previously identified key tasks in deliberation, we focused on assessing the potential management strategies considered and the dialogue or negotiation about these options. Three outcome variables characterized the deliberation: (1) the number of plans proposed, (2) the length of the deliberation, and (3) involvement of parent and/or child. Number of plans was a count of all unique plans proposed during the visit. The length of the deliberation was the longer of either the time between first and final plans or the time between first plan and final agreement. Deliberation length included time required to reach an agreement and periods of silent consideration of the options but specifically excluded interruptions that were not relevant to the deliberation. Involvement of the parent or the child was characterized as either active (child or parent proposed plans or disagreed with plans that were proposed by the physician) or passive (only the physician proposed plans).

Explanatory variables included child, physician, and parent characteristics as well as visit characteristics. Child characteristics included age, gender, and race/ethnicity (white non-Hispanic versus all others). Physician characteristics included gender, specialty (pediatrics or family practice), and years in practice. Parent characteristics included age, gender (mother/father/both parents/no parent in examination room), and education (college graduate/non–college graduate) of the accompanying parent(s). Visit characteristics included number of physician visits by the child annually (assessed as number of physician visits in the past 12 months) and number of visits to the participating physician by the child annually (assessed as number of visits to the participating physician in the past 12 months), an indicator of whether the physician was the child’s primary care physician and visit length. In the rare instance of >1 parent or >1 physician present, demographics of the participants who spoke more were used.

Analyses
Descriptive data included means, medians, SDs, and proportions. Interrater reliabilities were calculated with Cohen’s κ for ordinal or categorical variables and with intraclass correlation coefficients (ICCs) for continuous variables. Construct validity of the deliberation measures was evaluated with Spearman correlations. Because the number of plans proposed is count data with overdispersion, ordinary least squares regression is not appropriate for the analysis. Therefore, negative binomial models were used to evaluate the relationship between explanatory variables and the number of plans proposed. Incidence rate ratios (IRR) and 95% confidence interval (CIs) were computed for negative binomial regressions. IRRs describe the ratio of the number of plans proposed compared with the reference group or associated with a 1-unit change in the explanatory variable. For example, an IRR of 2.0 for girls indicates that girls proposed twice as many plans as boys. Gamma log generalized linear modeling was used to relate length of deliberation to participant and visit characteristics with deliberation time ratios and 95% CIs. Again, using child gender as an example, a deliberation time ratio of 2.0 for girls indicates that deliberation was twice as long with girls compared with boys. Logistic regression provided odds ratios (ORs) and 95% CIs for the associations between involvement and the explanatory variables. All models accounted for clustering within physician using robust estimates of variance.

We evaluated the following explanatory variables: child characteristics (age, gender, and race/ethnicity), physician characteristics (gender, specialty, and years in practice), parent characteristics (gender of the accompanying parent, age, and education), and visit characteristics (visit length, whether the physician was the child’s primary physician, and number of physician visits by the child annually). Parent age and number of visits to the participating physician by the child annually were not included in models because of collinearity with other model variables. Physician race/ethnicity was not included because the variable failed to achieve P < .20 in any model. Variables were modeled as either ordinal or continuous unless otherwise specified. All analyses were performed in Stata 8.

RESULTS
Participant Characteristics
Children’s mean age was 5.4 years (SD: 4.9; range: 0–17.5 years; Table 1). Forty-nine percent were female, and 75% were white non-Hispanic. Physicians were 61% female and 56% pediatricians, with a wide range of practice experience (years in practice mean: 13.0; SD: 8.3). Parents who accompanied the children were predominantly mothers, and 28% were college graduates. The visits averaged 12.2 minutes (SD: 5.3) in length, and 60% were to primary physicians. Eighty percent of the visits were for upper respiratory symptoms, with the remainder including other acute complaints such as gastrointestinal symptoms or injuries.

Reliability of Measures
Interrater agreement was almost perfect for speaker of plan (κ = 0.88) and was substantial for recipient of plan (κ = 0.63) and for number of disagreements (κ = 0.68). ICCs indicated substantial agreement for both the number of plans proposed (ICC: 0.70) and length of the deliberation (ICC: 0.70).
Number of Plans Proposed and Length of Deliberation
The mean number of plans proposed during a visit was 4.1 (median: 3.0). On average, approximately one quarter (2.9 minutes) of the visit was spent in deliberation. The proportion of the visit that was used for deliberation displayed wide variability (Fig 1). Four visits had no time spent in deliberation, whereas 9 visits had 50% of the visit time spent in deliberation. The majority of visits had less than one quarter of the time spent on deliberation.

Sharing of Deliberation
On average, physicians proposed 89% of plans during a visit, parents proposed 9% of plans, and children proposed 2% of plans. Of the plans that were proposed by the physician, 79% were proposed to the parent, 14% were proposed to the child, and 7% were proposed to the child and the parent jointly. Passive involvement of parents and children was observed in 65% of visits. The parent disagreed with a physician-proposed plan in 9% of visits. In all but 1 instance, the physician with whom the parent disagreed was female (P < .001).

Construct Validity of Deliberation Measures
We expected that the number of plans proposed and the length of deliberation would be highly correlated, indicating more discussion of potential plans. In addition, we expected that active parent/child involvement in deliberation would be positively correlated with the amount of deliberation. As expected, the number of plans proposed and the length of deliberation were highly correlated (ρ = 0.70, P < .001). Furthermore, the number of disagreements that were voiced in deliberation was positively correlated to both the number of plans proposed (ρ = 0.37, P < .001) and the length of deliberation (ρ = 0.34, P < .001). Also as expected, active involvement of parents and children was positively correlated with both the number of plans proposed (ρ = 0.50, P < .001) and length of deliberation (ρ = 0.35, P < .001).

Children’s Contributions to Deliberation
Children contributed to deliberation by both proposing plans and disagreeing with proposed plans. In 4% of visits, children proposed at least 1 plan. Children disagreed with a plan that the physician proposed in 6% of visits. The youngest of these children was 5.5 years of age (mean age: 11.25 years; range: 5.5–16). The dialogue in Appendix 2 depicts an 8-year-old boy’s involvement

---

**TABLE 1** Participant and Visit Characteristics (n = 101)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Child</strong></td>
<td></td>
</tr>
<tr>
<td>Age, mean (SD), y</td>
<td>5.4 (4.9)</td>
</tr>
<tr>
<td>Age group, %a</td>
<td></td>
</tr>
<tr>
<td>Infant/toddler, 0–2 y</td>
<td>45</td>
</tr>
<tr>
<td>Preschooler, 3–4 y</td>
<td>12</td>
</tr>
<tr>
<td>Grade-schooler, 5–11 y</td>
<td>27</td>
</tr>
<tr>
<td>Adolescent, 12–17 y</td>
<td>17</td>
</tr>
<tr>
<td>Female, %</td>
<td>49</td>
</tr>
<tr>
<td>White non-Hispanic, %</td>
<td>75</td>
</tr>
<tr>
<td><strong>Physician</strong></td>
<td></td>
</tr>
<tr>
<td>Female, %</td>
<td>61</td>
</tr>
<tr>
<td>Pediatrician, %</td>
<td>56</td>
</tr>
<tr>
<td>Family practice, %</td>
<td>43</td>
</tr>
<tr>
<td>Years in practice, mean (SD)</td>
<td>13.0 (8.3)</td>
</tr>
<tr>
<td><strong>Parent</strong></td>
<td></td>
</tr>
<tr>
<td>Age, mean (SD), y</td>
<td>34.5 (7.3)</td>
</tr>
<tr>
<td>Parent at visit, %</td>
<td></td>
</tr>
<tr>
<td>Mother only</td>
<td>78</td>
</tr>
<tr>
<td>Father only</td>
<td>11</td>
</tr>
<tr>
<td>Both parents</td>
<td>7</td>
</tr>
<tr>
<td>No parent in examination room</td>
<td>4</td>
</tr>
<tr>
<td>College graduate, %</td>
<td>28</td>
</tr>
<tr>
<td><strong>Visit</strong></td>
<td></td>
</tr>
<tr>
<td>Visit length, mean (SD), min</td>
<td>12.2 (5.3)</td>
</tr>
<tr>
<td>Visiting child’s primary physician, %</td>
<td>60</td>
</tr>
<tr>
<td>Annual physician visits, mean (SD)</td>
<td>6.3 (9.1)</td>
</tr>
<tr>
<td>Annual visits to participating physician, mean (SD)</td>
<td>3.2 (8.8)</td>
</tr>
</tbody>
</table>

*a Values do not add to 100% because of rounding.

---

**FIGURE 1** Distribution of percentage of visit spent in deliberation.
in the deliberation, resulting in the prescribing of a broader spectrum and more expensive second-line antibiotic, not the first-line agent recommended in clinical practice guidelines.61

**Deliberation and Child, Physician, Parent, and Visit Characteristics**

**Number of Plans Proposed**

In adjusted models, the number of plans proposed was related to visit characteristics and to child characteristics. Visit length was significantly associated with more plans proposed (IRR: 1.05; 95% CI: 1.03–1.08) in adjusted models (Table 2). Also, more plans were proposed when the child was female (IRR: 1.39; 95% CI: 1.08–1.78). No physician or parent characteristics were associated with the number of plans proposed.

**Length of Deliberation**

As with number of plans proposed, the length of the deliberation was related to characteristics of the visit but also to parent characteristics (Table 3). Shorter deliberations were found with college-graduate parents (deliberation time ratio: 0.62; 95% CI: 0.42–0.93). Longer deliberations occurred with longer visits (deliberation time ratio: 1.12; 95% CI: 1.07–1.17). No physician or child characteristics were associated with the length of deliberation.

---

**TABLE 2** Adjusted IRRs and 95% CIs for Factors Associated With Number of Plans Proposed (n = 101)

<table>
<thead>
<tr>
<th>Factor</th>
<th>IRR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, y</td>
<td>1.01</td>
<td>0.97–1.06</td>
</tr>
<tr>
<td>Female</td>
<td>1.39*</td>
<td>1.08–1.78*</td>
</tr>
<tr>
<td>White non-Hispanic</td>
<td>1.22</td>
<td>0.87–1.69</td>
</tr>
<tr>
<td>Physician</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.86</td>
<td>0.58–1.27</td>
</tr>
<tr>
<td>Pediatric</td>
<td>1.04</td>
<td>0.71–1.52</td>
</tr>
<tr>
<td>Years in practice</td>
<td>1.02</td>
<td>0.998–1.04</td>
</tr>
<tr>
<td>Parent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent at visit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother only</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Father only</td>
<td>1.05</td>
<td>0.87–1.28</td>
</tr>
<tr>
<td>Both parents</td>
<td>0.95</td>
<td>0.53–1.71</td>
</tr>
<tr>
<td>No parent in examination room</td>
<td>1.04</td>
<td>0.39–2.76</td>
</tr>
<tr>
<td>College graduate</td>
<td>0.85</td>
<td>0.66–1.11</td>
</tr>
<tr>
<td>Visit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visit length, min</td>
<td>1.05*</td>
<td>1.03–1.08*</td>
</tr>
<tr>
<td>Visiting child’s primary physician</td>
<td>0.93</td>
<td>0.62–1.39</td>
</tr>
<tr>
<td>Annual physician visits</td>
<td>0.99</td>
<td>0.96–1.01</td>
</tr>
</tbody>
</table>

*Adjusted for child characteristics (age, gender, and race/ethnicity), physician characteristics (gender, specialty, and years in practice), parent characteristics (gender of the accompanying parent and college graduate), and visit characteristics (visit length, whether the physician was the child’s primary physician, and number of physician visits by the child annually). IRRs describe the ratio of number of plans relative to the reference group or associated with a 1-unit change in the explanatory variable.  

---

**TABLE 3** Adjusted Deliberation Time Ratios and 95% CIs for Factors Associated With Length of Deliberation (n = 101)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Deliberation Time Ratio</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, y</td>
<td>1.02</td>
<td>0.98–1.06</td>
</tr>
<tr>
<td>Female</td>
<td>1.51</td>
<td>0.90–2.51</td>
</tr>
<tr>
<td>White non-Hispanic</td>
<td>1.05</td>
<td>0.60–1.82</td>
</tr>
<tr>
<td>Physician</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1.04</td>
<td>0.67–1.61</td>
</tr>
<tr>
<td>Pediatric</td>
<td>0.91</td>
<td>0.53–1.58</td>
</tr>
<tr>
<td>Years in practice</td>
<td>1.00</td>
<td>0.97–1.02</td>
</tr>
<tr>
<td>Parent</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent at visit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother only</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Father only</td>
<td>1.43</td>
<td>0.94–2.16</td>
</tr>
<tr>
<td>Both parents</td>
<td>0.73</td>
<td>0.27–1.97</td>
</tr>
<tr>
<td>No parent in examination room</td>
<td>1.70</td>
<td>0.64–4.56</td>
</tr>
<tr>
<td>College graduate</td>
<td>0.62*</td>
<td>0.42–0.93*</td>
</tr>
<tr>
<td>Visit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visit length, min</td>
<td>1.12*</td>
<td>1.07–1.17*</td>
</tr>
<tr>
<td>Visiting child’s primary physician</td>
<td>1.32</td>
<td>0.83–2.10</td>
</tr>
<tr>
<td>Annual physician visits</td>
<td>1.00</td>
<td>0.94–1.06</td>
</tr>
</tbody>
</table>

*Adjusted for child characteristics (age, gender, and race/ethnicity), physician characteristics (gender, specialty, and years in practice), parent characteristics (gender of the accompanying parent and college graduate), and visit characteristics (visit length, whether the physician was the child’s primary physician, and number of physician visits by the child annually). Deliberation time ratios represent the ratio of deliberation time relative to the reference group or associated with a 1-unit change in the explanatory variable.  

---

**Parent and Child Involvement in Deliberation**

In adjusted models, involvement of parents and children in the deliberation was associated with characteristics of the physician, parent, and visit (Table 4). Passive involvement was less likely with female physicians (OR: 0.22; 95% CI: 0.10–0.50) and with physicians with more years in practice (OR: 0.93; 95% CI: 0.88–0.98). Passive involvement was also less likely in longer visits (OR: 0.86; 95% CI: 0.79–0.94) and in visits when both parents were present (OR: 0.13; 95% CI: 0.05–0.37) compared with when the mother alone was present. The likelihood of passive involvement increased with the number of physician visits in the past year (OR: 1.11; 95% CI: 1.01–1.22). Child characteristics were not significantly associated with involvement in deliberation.

**DISCUSSION**

We developed and applied a reliable and valid technique to characterize the deliberative process in pediatric primary care visits, providing the first descriptions of how children, physicians, and parents deliberate together. These techniques complement existing tools that assess decision-making primarily from the standpoint of information exchange25–30 and extend previous work by evaluating deliberations that involve a patient and a caregiver. Although, on average, a substantial proportion of each visit was devoted to deliberation and the deliberation included the consideration of multiple plans, the
Physician 

Child

listed in information exchange during prescribing and making process across visits. Substantial variability exposure. In our study, we also noted considerable variation been noted when assessing the occurrence of information exchange during diagnostic and management plans, because fewer elements of informed consent during deliberation. This may be the result of our capturing deliberation for a 1-unit increase in the explanatory variable. For explanatory variables that continuous explanatory variables, OR represents the odds of passive parent/child involvement in deliberation when the characteristic is present.

Adjusted for child characteristics (age, gender, and race/ethnicity), physician characteristics (gender, specialty, and years in practice), parent characteristics (gender of the accompanying parent and college graduate), and visit characteristics (visit length, whether the physician was the child’s primary physician, and number of physician visits by the child annually). For continuous explanatory variables, OR represents the odds of passive parent/child involvement in deliberation for a 1-unit increase in the explanatory variable. For explanatory variables that reference the presence or absence of a characteristic, OR represents the odds of passive parent/child involvement in deliberation when the characteristic is present.

P < .05.

majority of visits exhibited passive involvement of parents and children. We also found specific child, physician, parent, and visit factors that were associated with the length of deliberation, the number of plans proposed, and parent/child involvement in deliberation. Given the influence of participation in decision-making for many common pediatric concerns, our work informs efforts to target and evaluate interventions to promote participation in the deliberative aspect of decision-making, particularly when such participation could improve outcomes.

Although previous studies of decision-making focused primarily on information exchange and neglected the critical component of deliberation, these studies noted considerable variability in the decision-making process across visits. Substantial variability existed in information exchange during prescribing and in the presence of elements of informed consent during routine clinical decisions. Similarly, variability has been noted when assessing the occurrence of information sharing, exploration of feelings or ideas, and closure. In our study, we also noted considerable variation in deliberation. This may be the result of our capturing both diagnostic and management plans, because fewer elements of informed decision-making are present during discussions of test ordering as opposed to management decisions.

Participatory decision-making is seen as the preferred approach in reaching many health care decisions, including decisions that are common to routine pediatric care, yet similar to our findings in children’s visits, only 50% of adult visits for hypertension or diabetes demonstrated evidence of participatory decision-making. Given that primary care physicians may well have greater knowledge of the values, expectations, and preferences of their patients and families, less child or parent involvement in these settings may be acceptable. However, more participatory approaches may be desirable or practical in situations in which there is more risk or more uncertainty—either about the outcome or about parent or child expectations, values, and preferences. Our work will support future studies to evaluate how parents and children feel about their involvement in deliberation and the decisions that result. This is particularly important in primary care pediatrics, because parent preferences have an influential role in many routine pediatric issues, such as circumcision and the use of antibiotics for upper respiratory infections, and physician perceptions of parental preferences are often inaccurate. This has been especially well evidenced in studies of antibiotic use for upper respiratory symptoms.

As with studies of patients’ perceptions of participation, visit length was positively associated with the number of plans proposed, the length of deliberation, and more parent/child involvement in deliberation, all suggesting more participatory deliberation in longer visits. The Institute of Medicine cites reduced visit lengths as barriers to effective health care communication, with shorter visits associated with worse patient and physician outcomes. This may be particularly problematic in geriatric or pediatric visits, in which patients are commonly accompanied by a caregiver, necessitating the consideration of 3 individuals’ expectations, values, and preferences.

In our analysis, there was less deliberation in visits with male children and with college-graduate parents, whereas more parent/child involvement in deliberation was found in visits to female and more experienced physicians. The first finding regarding child gender is consistent with the adult literature: adult female patients experience more participatory visits. However, our result with regard to parent education is contrary to subjective assessments of participation among adult patients, where better educated patients report more participation. Our findings about parent/child involvement are consistent with patient reports of more participatory visits with female physicians and with previous work demonstrating that more experienced physicians were more likely to report involving parents in management decisions for infants who are at the border of viability. The latter result suggests that training residents to involve children and parents in deliberation may be necessary and fruitful, especially for pediatric primary care issues that are preference sensitive.
The small sample size for analysis may have limited our ability to detect significant associations, particularly for the physician factors. Despite this, we did find several plausible and consistent associations. For our initial exploration of deliberation, we limited the events that were considered to the proposing of plans and agreements or disagreements in response to those proposals. For provision of a finer assessment of the deliberative process, future work will include assessments of specific content elements in deliberation, such as discussions of risks, benefits, adverse effects, and uncertainty of outcomes for proposed plans. Furthermore, as in all nonrandomized, observational studies, unmeasured factors may contribute to associations observed. We did not adjust our results for the severity of child illness. However, we did exclude from participation any child who seemed acutely ill. Last, although we recruited a diverse group of healthy participants, caution in generalizing results is warranted, particularly when generalizing to the care of children with chronic illness.

CONCLUSIONS
To our knowledge, this work provides the first look at deliberation in pediatric visits. Although deliberation comprised a substantial portion of the acute pediatric visit and included many proposed plans, passive involvement of parents and children predominated. Given that children’s participation improves their health outcomes and the quality of the health care provided while being more satisfying for parents and children, striving for a more participatory approach seems advisable, and future studies should evaluate the benefits of greater participation in both information exchange and deliberation. Furthermore, because both parents and children feel ill-prepared to participate in health care decisions and adult patients can experience anxiety when encouraged to participate, more work is needed to support the development of participation skills during childhood. This may be particularly important in primary care pediatric visits, because decisions for many routine pediatrics issues can be influenced by participation and preferences. Our technique for characterizing participation in deliberation for the child, physician, and parent represents a significant advance in facilitating such research, complementing previous work that has focused on information exchange. Last, for pediatric primary care scenarios in which participation in decision-making can improve outcomes, our results inform interventions to promote participation by considering how child, physician, parent, and organizational factors influence participation in the deliberative process. If results are confirmed, advocating for policies that foster deliberation (eg, longer primary care visit lengths) may encourage child and parent participation and ultimately result in improved health outcomes.

APPENDIX 1: APPLICATION OF DELIBERATION CODING SCHEME TO VISIT DIALOGUE
Parent: “So what do you think, Amox?” [Plan Proposed]
Doctor: “No, no, he’s got this bad one (looking at medical chart) here.” [Disagree]
Parent: “I know.”
Doctor: “I think our 2 choices that we’ve got are Biaxin [Plan Proposed] and Vantin.” [Plan Proposed]
Parent: “Whichever. What we went through with his ear infections was Vantin, and last time that didn’t work. But, whatever you think. Then the next one we did, I think, was the Biaxin, and I think that worked.”
Doctor: “And he didn’t react to it . . . like the Zithromax?”
Parent: “Well, you know, the Zithromax was questionable anyway . . . . We don’t know if it was (the) medicine, or viral, or what.”
Doctor: “Does he take the Biaxin okay?”
Parent: “He took it fine.”
Doctor: “Okay. Why don’t we stick with Biaxin.” [Final Plan]
Parent: “Okay.” [Agreement]

APPENDIX 2: ILLUSTRATION OF 8-YEAR-OLD CHILD’S INVOLVEMENT IN DELIBERATION
Doctor: “You have a significant sinus infection, so we have to treat that. Are you allergic to any medicines?”
Child: “No, just vitamins!”
Doctor: “Just vitamins? . . . Do you chew pills, do you swallow pills, or do you drink?”
Mother: “He chews. Well, what would you rather do, chew or drink?”
Child: “I want to drink.”
Doctor: “Well, we have to give you 3 weeks worth of the medicine that works.”
Child: “No bubble gum medicine.”
Mother: “No bubble gum flavors.”
Doctor: “Well, then, we better not do liquid, because it’s bubble gum flavored.”
Child: “Yuck . . . .”
Mother: “Honey, you can have a drink of water right after. Otherwise, you’ll have to chew a pill.”
Child: “Okay, I’ll chew a pill!”
Mother: “You’re sure?”
Child: “What is it?”
Doctor: “It’s the same medicine, except it’s in a pill instead of a liquid.”
Child: “I’ll chew a pill.”
Doctor: “It might be bubble gum, I don’t know.”
Child: “Yuck!”
Mother: “Wouldn’t you just rather have the liquid and get it over with?”
Child: “I don’t like bubble gum. Cherry, Motrin, and berry, that’s what I like.”
Doctor: “We’ll give you a different kind that doesn’t taste
cherry or berry . . . this is called Augmentin, and it is not bubble gum flavored.”

ACKNOWLEDGMENTS
This work was funded by the Agency for Healthcare Research and Quality via K08-HS13183.

We gratefully acknowledge participation by All Saint Medical Group (Racine, WI); UW Health Clinics; Beloit Clinic; Lodi Clinic and Meade Clinic; coding of data by Saurabh Saluja, Sonja Raaum, and Tim Conrad; manuscript feedback from Drs Mark Linzer and Patricia Kokotailo as well as Access Quality and Outcomes Research Network group members; and manuscript preparation by Sonja Raaum, Matt Swedlund, and Julia Yahneke. Linda Sanders and Kirstin Muehlbauer trained and monitored coders; Ron Cramer prepared videotapes.

REFERENCES
21. Odigwe C. Children say they are not involved in their treatment. BMJ. 2004;328:600


56. The Observer Basic [computer program]. Version 5.0. Leesburg, VA: Noldus Information Technology; 2004

57. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174


60. *Stata Statistical Software* [computer program]. Version 8.0. College Station, TX: Stata Corp; 1999


69. Institute of Medicine. *Health Literacy: A Prescription to End Confusion.* Washington, DC: Committee on Health Literacy and Board on Neuroscience and Behavioral Health; 2004


80. Mounjnd J, Carrere MO, Charavel M, Bremond A. Clinical
issues in shared decision-making applied to breast cancer. *Health Expect*. 2003;6:222–227


ARTICLE

Prenatal Cocaine and Tobacco Effects on Children’s Language Trajectories

Barbara A. Lewis, PhD, H. Lester Kirchner, PhD, Elizabeth J. Short, PhD, Sonia Minnes, PhD, Paul Weishampel, MA, Sudtida Satayathum, MS, Lynn T. Singer, PhD

Departments of Pediatrics, Psychology, and General Medical Sciences, Case Western Reserve University, Cleveland, Ohio

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

ABSTRACT

OBJECTIVE. The objective of this study was to examine the effects of prenatal cocaine and polydrug exposure on language development of preschool children using a prospective longitudinal model, controlling for confounders.

METHODS. Children who were exposed to cocaine in utero (n = 209) and nonexposed children (n = 189) were followed prospectively at birth and at 1, 2, 4, and 6 years of age and were compared on receptive, expressive, and total language scores across time using random coefficient models, controlling for confounders.

RESULTS. A significant, stable effect of cocaine exposure on language development was observed over time for all language domains, with cocaine exposure related to poorer language performance. Cigarette exposure was related to lower receptive language scores. Environmental influences on language scores were also observed. Both the cocaine-exposed and nonexposed children declined in language performance over time.

CONCLUSIONS. Prenatal cocaine exposure has a stable negative effect on language skills during the first 6 years of life. Both cocaine-exposed and nonexposed children showed decreased language growth over time; however, cocaine-exposed children demonstrated linguistic deficits compared with nonexposed peers and did not catch up. Cigarette and environmental influences were also noted.
Children who are exposed to cocaine in utero are at risk for a variety of developmental delays as a result of both biological risk and postnatal environmental influences. Biological risks include a disruption in arousal and attention.1-5 The environmental risks are threefold: inadequate stimulation provided by a drug-using mother,6,7 insecure child attachment,6 and poverty.8 These risk factors may have an impact on the language development of children who are exposed to cocaine.

Outcome studies of language development of children who were exposed to cocaine in utero have been equivocal,4,10-20 possibly as a result of methodologic differences such as age at assessment, instruments used, and variability in selection of confounding factors considered.15 Although most studies have reported language outcomes at a single time point, one of the few longitudinal prospective studies of children who were exposed to cocaine found stable cocaine-specific effects on total language scores, even after control for multiple medical and demographic covariates.11,21 Children with cocaine exposure performed on average 15% of a SD lower on a standardized language test than nonexposed children, with the strongest effects at 18 months and 3 years. This study extended these findings with a larger developmental age span, examining also receptive and expressive language and using biological markers (meconium) of cocaine exposure.

This investigation examined the effects of cocaine exposure in utero on language development in a longitudinal sample of children who were enrolled prospectively at birth and followed to 1, 2, 4, and 6 years of age. Children who were exposed to cocaine in utero were hypothesized to perform more poorly than nonexposed children on standardized measures of receptive and expressive language across all time points. Careful delineation of environmental factors that are known to relate to child language skills was conducted, particularly maternal education/vocabulary, depressive symptoms, and use of other substances.

Methods

Subjects
A total of 398 children (209 cocaine exposed [CE] and 189 non–cocaine exposed [NCE]) were followed prospectively from birth and assessed for language development at 1, 2, 4, and 6 years of age. The sample was drawn from a cohort that was recruited at birth from a large, urban, county teaching hospital to participate in a longitudinal study of the sequela of fetal drug exposure. Approval from University Hospitals of Cleveland Institutional Review Board for Human Investigation and informed consent were obtained. Women who were considered at high risk for drug use were administered drug toxicology screenings. Urine samples were obtained immediately before or after labor and delivery and analyzed for the presence of cocaine metabolites (benzoylecgonine), cannabinoids, opiates, phencyclidine, and amphetamines. In addition, infants had meconium drug analyses performed for cocaine and its metabolites (benzoylecgonine [BZE], meta-hydroxybenzoylecgonine [M-OH-BZE], cocaethylene, cannabinoids, opiates, phencyclidine, amphetamines, and benzodiazepines). CE infants were identified on the basis of positive infant meconium, maternal urine, or maternal self-report, whereas control infants were negative on all indicators. Women who used alcohol, marijuana, or tobacco during pregnancy were included in both groups.

The sample size for the original cohort was 415. The number of children who participated in the language testing at each follow-up time point varied; however, ~85% of the cohort participated in 3 or more of the 4 visits. At 1 year, the sample size consisted of 405 children, with 371 assessed and 265 children receiving the language measures. At 2 years, 404 children from the original sample were available (1 death), with 381 assessed and 339 completing language measures. At 4 years, 404 children were available, with 394 completing the language test, 12 (8 CE) not coming to the visit, 16 (12 CE) dropouts, and 2 (1 CE) moving out of state. Attrition was greater for the CE group (P = .04) than for the NCE group. At 6 years of age, 377 received assessments, with 371 children completing the language test battery.

Procedures
At 1 and 2 years of age, the Preschool Language Scale, Third Edition22 was administered. At 4 years, the Clinical Evaluation of Language Fundamentals–Preschool23 was given. At 6 years, the Comprehensive Assessment of Spoken Language24 was administered. At all ages, examiners were unaware of infant cocaine status.

For assessment of prenatal drug exposure, infants and their biological mothers were seen immediately after birth, at which time the biological mother was interviewed regarding drug use. Biological mothers were asked to recall the frequency and the amount of drug use for the month before pregnancy and each trimester of her pregnancy. More specific, for tobacco, the number of cigarettes smoked per day was recorded; for marijuana, the number of joints smoked per day was recorded; for alcohol, the number of drinks of beer, wine, or hard liquor per day was computed; and for cocaine, the number of rocks consumed and amount of money spent per day were noted. This drug assessment was updated at each follow-up visit to provide a similar measure of current drug use, with the assessments also administered to the foster or relative caregiver to provide a measure of postnatal environmental exposure for children who were placed out of maternal care.

Birth, demographic, and medical characteristics were taken from hospital records and included maternal race,
age, parity, number of prenatal care visits, and type of medical insurance, infant Apgar scores, birth weight, length, and head circumference. At enrollment, maternal socioeconomic status (SES; A. B. Hollingshead, PhD, *Four-Factor Index of Social Status*, unpublished manual, 1975) and educational level were calculated. Maternal vocabulary score was measured using the Peabody Picture Vocabulary Test, Revised (PPVT-R). Two subtests of the Wechsler Adult Intelligence Scale, Revised^2^6 (WAIS-R) were administered: The Block Design and Picture Completion subtests from the WAIS-R enabled an estimation of nonverbal intelligence. The Brief Symptom Inventory^2^7 is a standardized self-report scale that was administered at birth and at all visits to obtain a measure of severity of psychological distress. The General Severity Index, a summary score of the Brief Symptom Inventory, was used as an indicator of overall distress. The Hobel Neonatal Risk Index^2^8 was computed to obtain a measure of neonatal medical complications. Also at the visit, the child’s placement (either biological mother/relative or foster/adoptive caregiver) was noted, and data on the current caregiver were updated. When the child had been placed with a new caregiver, intellectual measures of the caregiver were also updated. The Home Observation of the Environment (HOME), Preschool version was administered to the caregiver in an interview format as a measure of the quality of the caregiving environment.^2^9

**Statistical Analysis**

Baseline maternal and child characteristics and prenatal drug exposure were summarized using means and SDs for continuous variables and frequencies and percentiles for categorical variables. Comparisons between CE and NCE groups were performed using t tests, Wilcoxon rank sum tests, and Pearson χ² tests. All positively skewed data, including drug self-report measures and General Severity Index, were transformed using the natural logarithm of (x + 1) to achieve a distribution that approximates normality. Correlations between drug exposure data and language outcomes were estimated using Spearman correlation coefficients.

For examination of language performance across time, each measurement was internally standardized to create z scores at each visit using all available children. Standardization of scores at each time point allowed modeling to be performed across language measures despite that different age-appropriate tests were used. Analyses of the z scores were accomplished using random coefficient models with restricted maximum likelihood estimation. The intercept and the slope for child age were treated as random effects to capture the variability and correlation in the data. The actual age of the child was used instead of visit age to capture trends over time better. These models were used to estimate and test relationships between cocaine exposure groups at the follow-up visits (ages 1, 2, 4, and 6 years). Initially, child age, cocaine exposure, and the interaction between age and exposure were included in the model to test for changing effects over time. Because the interaction term was not significant for any of the language measures, main effects models were fit. The developmental trajectories were allowed to be nonlinear (eg, quadratic) using polynomials of time. The lowest order polynomial of time was retained. In all models, the trajectories were found to be linear. The effects of cocaine are presented with and without consideration of possible confounding variables.

For each measure, the model-building strategy of Bandstra et al^1^1 was followed to achieve a final model that controlled for possible confounding and moderating variables to estimate an unbiased effect of prenatal cocaine exposure on language development. The following variables were considered for each outcome: child’s age, prenatal cocaine exposure (yes versus no), and the interaction between age and cocaine exposure; maternal age at child’s birth, current caregiver’s PPVT-R and WAIS-R block design; child’s race and child’s gender; prenatal drug variables (alcohol, cigarettes, and marijuana), prenatal care, parity, SES, and marital status; current HOME scale; and adoptive/foster care. Adjusted least squares means and SEs were calculated from the final models and compared between treatment groups at each follow-up visit age. Plots of the adjusted group means (±SE) for each language measure over time are provided. Analyses were performed using SAS 9 (SAS Institute, Cary, NC).

**RESULTS**

**Sample Characteristics**

Cocaine-using women and control subjects were primarily black, of low income, and not married (Table 1). Cocaine-using women were older, had more children, and attended fewer prenatal care visits. They used other drugs more frequently and in higher amounts than nonusers (Table 2). CE infants were more likely to be preterm and of lower birth weight, head circumference, and birth length than NCE infants (Table 3).

At birth, 49 (26%) CE infants were placed outside maternal/biological care compared with only 3 (2%) of NCE infants. By 4 years, 42 (22%) CE children were in adoptive/foster care compared with 10 (8%) of NCE children. CE children averaged 1.0 ± 0.99 nonmaternal care placements by 6 years versus 0.16 ± 0.48 for NCE infants (*t* = −10.12; *P* < .0001). There were no group differences in HOME environment scores.

From birth to 6 years, there were 11 deaths (8 CE and 3 NCE; χ² = 1.9, *P* < .17). Causes of death for the CE children were sudden infant death syndrome (*n* = 4), cardiopulmonary arrest (*n* = 1), pneumonia (*n* = 1), accidental asphyxia (*n* = 1), and respiratory distress syndrome (*n* = 1). Causes of death for the NCE children...
were sudden infant death syndrome (n = 2) and respiratory distress syndrome (n = 1). From enrollment at birth, the retention rate was 93% (377) at 6 years for surviving children.

Meconium Assays and Outcomes

Several significant relationships were found between the concentration (ng/g) of cocaine metabolites and child language outcomes. At 1 year, the concentrations of BZE (r = −0.16, P < .02) and M-OH-BZE (r = −0.14, P < .04) were negatively related to the expressive language score; and at 2 years, the concentration of BZE was negatively related to receptive (r = −0.14, P < .02), expressive (r = −0.12, P < .05), and total language scores (r = −0.11, P < .02). At 4 years, the concentration of cocaethylene, the metabolite formed through the

---

TABLE 1  
Maternal Demographics for CE and NCE Groups

<table>
<thead>
<tr>
<th>Maternal Demographics</th>
<th>CE Group (N = 209)</th>
<th>NCE Group (N = 189)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at birth</td>
<td>29.68</td>
<td>25.71</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Years of education</td>
<td>11.57</td>
<td>11.99</td>
<td>.07</td>
</tr>
<tr>
<td>Parity</td>
<td>3.53</td>
<td>2.71</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>No. of prenatal visits</td>
<td>5.23</td>
<td>8.78</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>PPVT-R</td>
<td>73.78</td>
<td>77.82</td>
<td>.09</td>
</tr>
<tr>
<td>WAIS-R Design Scale</td>
<td>6.86</td>
<td>7.15</td>
<td>.20</td>
</tr>
<tr>
<td>WAIS-R Picture Completion</td>
<td>6.60</td>
<td>7.02</td>
<td>.07</td>
</tr>
<tr>
<td>Global Severity Index</td>
<td>0.82</td>
<td>0.49</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Black, n (%)</td>
<td>169 (82.04)</td>
<td>153 (80.95)</td>
<td>.78</td>
</tr>
<tr>
<td>Currently employed, n (%)</td>
<td>11 (5.37)</td>
<td>40 (21.28)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Married, n (%)</td>
<td>16 (7.77)</td>
<td>32 (16.93)</td>
<td>.05</td>
</tr>
<tr>
<td>Low SES, n (%)</td>
<td>201 (98.05)</td>
<td>185 (97.88)</td>
<td>.91</td>
</tr>
</tbody>
</table>

TABLE 2  
Maternal Drug Use During Pregnancy for CE and NCE Groups

<table>
<thead>
<tr>
<th>Maternal Drug Use During Pregnancy</th>
<th>CE Group (N = 209)</th>
<th>NCE Group (N = 189)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tobacco, cigarettes per d</td>
<td>11.57</td>
<td>4.11</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Alcohol, dosage per wk</td>
<td>10.16</td>
<td>1.35</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Marijuana, dosage per wk</td>
<td>1.36</td>
<td>0.59</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Cocaine, units per wk</td>
<td>23.56</td>
<td>4.42</td>
<td>.0011</td>
</tr>
<tr>
<td>Prevalence of Use, n (%)</td>
<td>Tobacco</td>
<td>174 (87.88)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Alcohol</td>
<td>171 (86.36)</td>
<td>119 (65.75)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Marijuana</td>
<td>96 (47.33)</td>
<td>23 (12.71)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Amphetamine</td>
<td>5 (2.54)</td>
<td>2 (1.11)</td>
<td>.31</td>
</tr>
<tr>
<td>Barbiturate</td>
<td>1 (0.51)</td>
<td>1 (0.55)</td>
<td>.95</td>
</tr>
<tr>
<td>Heroin</td>
<td>5 (2.55)</td>
<td>0 (0)</td>
<td>.03</td>
</tr>
<tr>
<td>Phencyclidine</td>
<td>11 (5.64)</td>
<td>0 (0)</td>
<td>.01</td>
</tr>
</tbody>
</table>

TABLE 3  
Child Demographics for CE and NCE Groups

<table>
<thead>
<tr>
<th>Child Demographics</th>
<th>CE Group (N = 206)</th>
<th>NCE Group (N = 189)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Mean</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight, ga</td>
<td>2726</td>
<td>3100</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Birth length, cm²</td>
<td>47.36</td>
<td>49.10</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Head circumference, cm²</td>
<td>32.31</td>
<td>33.46</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Microcephaly</td>
<td>30</td>
<td>10</td>
<td>.002</td>
</tr>
<tr>
<td>Apgar score (1 min)</td>
<td>8.00</td>
<td>7.93</td>
<td>.61</td>
</tr>
<tr>
<td>Apgar score (5 min)</td>
<td>8.79</td>
<td>8.78</td>
<td>.89</td>
</tr>
<tr>
<td>HOBEL Neonatal Risk score</td>
<td>7.27</td>
<td>5.67</td>
<td>.32</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>94 (45.63)</td>
<td>92 (48.68)</td>
<td>.54</td>
</tr>
<tr>
<td>Black, n (%)</td>
<td>168 (81.55)</td>
<td>152 (80.42)</td>
<td>.77</td>
</tr>
<tr>
<td>Preterm (&lt;37 wk gestational age), n (%)</td>
<td>58 (28.16)</td>
<td>35 (18.52)</td>
<td>.02</td>
</tr>
<tr>
<td>Low birth weight (&lt;2500 g), n (%)</td>
<td>72 (34.95)</td>
<td>34 (17.99)</td>
<td>.0001</td>
</tr>
<tr>
<td>Very low birth weight (&lt;1500 g), n (%)</td>
<td>12 (5.83)</td>
<td>7 (3.70)</td>
<td>.33</td>
</tr>
<tr>
<td>Small for gestational age, n (%)</td>
<td>25 (12.32)</td>
<td>4 (2.13)</td>
<td>.0001</td>
</tr>
</tbody>
</table>

* Adjustment for gestational age to these measures.
combining cocaine and alcohol, was negatively related to expressive language \( (r = -0.12, P < .03) \) and marginally related to total language score \( (r = -0.11, P < .06) \). By 6 years of age, there were no significant relationships.

**Longitudinal Analyses**

Table 4 presents the mean unadjusted receptive, expressive, and total language scores by group for each of the 4 time points. Table 5 presents the adjusted mean language scores at each visit for the CE and NCE groups.

**Total Language**

The interaction between child age and cocaine exposure was not significant, suggesting that total language development demonstrated a similar growth trajectory across the groups. In the final model, a main effect for race was found, indicating that black children scored almost 0.5 of a SD lower than nonblack children. Significant interactions of child age with maternal vocabulary score, maternal age, and child gender were found. Gender was associated with total language scores at 1, 2, and 4 years of age, with boys performing more poorly. Maternal vocabulary was significant at 4 and 6 years, with higher maternal vocabulary related to higher child total language scores. Older maternal age at the child’s birth was significant only at 6 years, with older maternal age related to poorer language scores. Controlling for these covariates, a borderline significant effect of cocaine exposure was found \( (P = .058) \). CE children performed \( -0.15 \) SD \( (±0.08) \) lower than NCE children. No effects of marijuana or alcohol exposure were found. However, cigarette exposure was negatively associated with receptive language \( (P = .0168) \). Marital status and adoptive/foster care were not significantly related to receptive language.

**Expressive Language**

The interaction between child age and cocaine exposure was not significant for expressive language scores. Significant main effects for race and gender were found. Black children scored approximately one third of a SD lower than nonblack children, and boys performed one third of a SD lower than girls. The expressive language trajectory also varied by current HOME score \( (P = .0002) \). Higher HOME scores were associated with higher expressive language scores at 2 and 4 years of age \( (P < .0001) \). Controlling for these covariates, a significant and constant effect of cocaine exposure was found \( (P = .484) \). CE children perform \( -0.15 \) SD \( (±0.08) \) lower than NCE children. Marijuana, alcohol, or cigarette exposure was not a significant confounder. Marital status and adoptive/foster care were not related to expressive language development. On average, all children declined in their expressive language scores at 4 years compared with the 1-year visit \( (P = .03) \).

**Associations of the PPVT-R, HOME, and Cigarette Exposure to Language**

The effect of current caregivers’ PPVT-R was significant at 4 years of age, such that the PPVT-R scores increased, so did the child’s scores on receptive \( (P = .0066) \) and total language \( (P = .0025) \) measures. This effect was also observed for the total language score at 6 years of age \( (P < .001) \). The HOME scale showed significant effects across the ages. At 1 year of age, the receptive language score was negatively related to the HOME scale, such that as the HOME score increased, the receptive language score decreased \( (P = .04) \). At 2 years of age, the HOME score was positively related to expressive language, such that as the HOME score increased, so did the child’s expressive language score \( (P = .037) \). At 4 years of age, the HOME score was positively related to all language scores (receptive, \( P < .0001 \); expressive, \( P < .0001 \); and total, \( P = .005 \) scores). This effect remained at 6 years of age \( (P = .001) \).

Children who were exposed to any cigarette smoking...
in utero had a lower mean standardized receptive language score by 0.21 ± 0.09 SD (95% confidence interval: −0.38 to −0.04). Children who were exposed to cigarette smoking also had a lower mean standardized total language score by 0.17 ± 0.09 SD (95% confidence interval: −0.35 to 0.01). Exposure to cigarette smoking did not have a significant effect on expressive language at any time point and was therefore not included in the final model for expressive language.

**DISCUSSION**

This study presents a stable picture of the negative effects of prenatal cocaine exposure on language skills during the first 6 years of life. Correlations of the concentration of cocaine metabolites in infant meconium with later language outcomes also indicated that cocaine exposure affects language development. CE children had poorer language scores at all 4 ages than the NCE children. Although the CE group showed stable language growth over time, they did not catch up to their NCE peers. These findings are consistent with and extend the findings of Bandstra et al., who also found a stable cocaine-specific effect on total language scores between the ages of 3 and 7 years. We examined receptive and expressive language scores at 3 time points and found marginally significant effects of cocaine exposure over time. These findings concur with previous reports of both receptive and expressive language delays in children who were exposed to cocaine.

Of note is that language scores in both the CE and NCE groups declined over time, suggesting that factors that are common to both groups, such as low SES, education, and poverty, have a negative impact on the developmental trajectory of language, a finding that is in agreement with the longitudinal studies of the Miami Prenatal Cocaine study that also report a decline in scores on standard language measures with time. Cigarette smoking during pregnancy also had a negative impact on receptive language skills.

These findings are consistent with previous research by Fried and Watkinson that showed reduced auditory processing skills in children who were exposed prenatally to tobacco. In the neonatal period, those infants demonstrated decreased rates of auditory habituation.

At 12 to 24 months, infants showed poorer responses to auditory-related items on the Bayley Scales of Infant Development. Follow-up at 3 and 4 years of age revealed deficits in language skills and at 6 years in auditory processing skills. At 9 to 12 years, these children presented with lower language and reading scores, particularly related to the auditory aspects of these skills. The present findings indicate...
that tobacco exposure is additive to the risk of CE children for language deficits.

**Environmental Effects**

Our data also support the notion that environmental variables can affect language skills to a considerable extent. That current caregiver’s vocabulary score and the HOME score both seem to have an impact on language performance underscores the environmental modifiability of language. Bandstra et al. found a relationship between language outcomes and the HOME scale. Previous reports of delay in semantic representation in children who were exposed to cocaine may relate to the caregiver’s vocabulary rather than to the cocaine exposure itself. The current caregiver’s vocabulary score was significant only at the 4-year and 6-year testing times but not at the younger ages. At a young age, the child’s language development may not be as influenced by the caregiver’s vocabulary because the child is acquiring basic vocabulary and syntactic structures. However, as language development proceeds, the child is required to master complex syntax, and vocabulary growth is extremely rapid. That a caregiver’s verbal skills played an important role in the child’s language growth is interesting. As reported previously, the foster/adoptive caregiver’s vocabulary, depression, and HOME scores mediate the adoptive-care effect on language at 4 years of age.

**Limitations of Study**

Several limitations of this study should be noted. First, not all children were assessed at each time point, with the fewest number of children completing the language test at 1 year of age. However, 85% of the enrolled sample are represented in at least 3 time points. A second limitation was the use of different language measures at different time points. The preschool years are a time of rapid language acquisition, and few measures can adequately assess language skills from 1 year to 6 years of age. Ceiling and floor effects are problematic. Despite that different assessments were used at each time point, the magnitude of language deficit between the CE and NCE children remained constant at each time. This finding suggests that the different measures were assessing similar language constructs.

**CONCLUSIONS**

The findings from this study support the notion that language outcomes are the result of both drug exposures and environmental factors. The cumulative risk for language disorders is likely to be based on prenatal drug or other toxic exposure, environmental and genetic influences, and social factors. Future studies should attempt to identify specific linguistic deficits in semantic, syntactic, phonologic, and pragmatic skills that are associated with cocaine and tobacco exposures and determine whether specific deficits change over time. Pediatricians should be aware of the additional risks that prenatal cocaine and tobacco exposures have on language development of poor, urban children and increase their surveillance and intervention efforts in these populations.

**ACKNOWLEDGMENTS**

This study was supported by grants R01-07259 and R01-07957 from the National Institute on Drug Abuse.

We thank Terri Lotz-Ganley for manuscript preparation; Kristen Weigand, Laurie Ellison, Selena Cook, Julia Noland, Martin Manuel, Astrida Seja Kaugars, and Teresa Linares for research and data analytic assistance; and the Cuyahoga County Department of Children and Family Services.

**REFERENCES**

15. Lewis BA, Singer LT, Short EJ, et al. Four-year language out-
comes of children exposed to cocaine in utero. *Neurotoxicol Teratol.* 2004;26:617–627
Self-Reported Health Status and Health-Related Quality of Life of Teenagers Who Were Born Before 29 Weeks’ Gestational Age

Ron Gray, MBChB, MPH*, Stavros Petrou, PhD, MPhil*, Christine Hockley, BA*, Frances Gardner, DPhilb

*National Perinatal Epidemiology Unit and bDepartment of Social Policy and Social Work, University of Oxford, Oxford, United Kingdom

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

ABSTRACT

OBJECTIVE. The objective of this study was to describe the self-reported health status and health-related quality of life of British teenagers who are in mainstream schooling and were born before 29 weeks’ gestational age compared with British teenagers who were born at term.

METHODS. All surviving children who were born at <29 weeks’ gestation in the former Northern Region of England in 1983 and in the former Oxford Region of England and in Scotland in 1984 were eligible. A comparison group of teenagers who were born at term were also recruited. Children’s responses to the Health Utilities Index Mark III were compared.

RESULTS. A total of 218 of the original 535 children who were born in the 3 regions during the study period were alive at 15 to 16 years of age. A complete Health Utilities Index Mark III record was available for 140 children in mainstream schools and for 108 control subjects. In 7 of the 8 attributes (vision, hearing, speech, emotion, pain, ambulation, and dexterity), there were no statistically significant differences in any functional impairment between the comparator groups. However, the preterm group did report a higher level of functional impairment in the cognition attribute (40.7% vs 25.0%). Although there was no difference in the median Health Utilities Index Mark III utility score between the 2 groups (0.93), there was a broader range of utility scores for the preterm group (0.07–1.0 vs 0.45–1.0 for the control group).

CONCLUSIONS. Despite objective evidence that children and teenagers who were born preterm have poorer health on average than term-born control subjects, this is not reflected in their own ratings of their health status and health-related quality of life at 15 to 16 years of age. The reasons for these differences need to be further explored.
Children who are born very preterm are at increased risk for a range of adverse neonatal outcomes, including chronic lung disease, severe brain injury, retinopathy of prematurity, necrotizing enterocolitis, and neonatal sepsis. In later life, children who were born very preterm are at increased risk for motor and sensory impairment learning difficulties, and behavioral problems. Most studies that describe the long-term outcomes of children who are born very preterm use disease-specific instruments that fail to capture all neurodevelopmental, functional, and behavioral outcomes that might be of interest. In recent years, a number of investigators have recognized the importance of measuring the impact of preterm birth across multiple domains. Instruments that can be used to measure the multiple health impacts of preterm birth over the longer term include multidimensional health profiles, which measure different aspects of physical, mental, and social well-being, and multiattribute utility measures, which are health status classification systems with preexisting preference weights that can be attached to each permutation of responses. A particular advantage of the latter set of measures is that they generate composite utility scores that reflect population preferences for the overall health state that is being measured.

The multiattribute utility measures that have been developed include the Quality of Well-Being Scale, Rosser-Kind Classification of Illness States, Health Utilities Index (HUI), EuroQol 5-dimension, 16D (a 16-dimensional measure of health-related quality of life), 17HUI Mark II classification. Differences between the estimates, the ultrasound estimate was available. When there was a discrepancy of >14 days between these estimates, the ultrasound estimate was preferred.

For locating the children at age 15 to 16 years, a letter was sent to the address of the general practitioner who was identified from the child’s most recent follow-up assessment. When the child was still registered, permission was requested from the general practitioner to reapproach the family. When the child was no longer registered with that general practitioner, the National Health Service Central Register was used to confirm that the child was still alive and, if so, the location of the child’s current general practitioner. The child’s current general practitioner was then approached and asked permission to reapproach the family.

Recruitment procedures varied among the 3 collaborating regions of the ELGA study, but essentially both the child and the parents were asked to give written permission for questionnaires to be sent to the child.
parent, and general practitioner and for the child's school to be approached. The head teacher for each consenting child in mainstream schooling was asked to forward a questionnaire to the child's year tutor for completion and to help locate children who could act as study control subjects. The year tutor was asked to choose as control subjects the 3 children who were closest in date of birth to the ELGA child, in the same year group, and of the same gender. Control children and their parents were then written to and asked permission to participate in the study. Control subjects were not sought for children who were attending special schools. Full details on the tracing and recruitment procedures of the study participants are reported elsewhere.34

Measurement of Health Status and Health-Related Quality of Life

The postal questionnaires that were sent to the children in 1999–2000 included questions on their school progress, health, mental health, behavior problems, and substance misuse. The children were asked to complete the postal questionnaires without assistance from their parents or other caregivers. Because the focus of the analysis was health status and health-related quality of life, we report in this article their responses to the HUI questionnaires. The HUI is a family of health status classification systems. The children were asked to complete the unedited 15-item questionnaire for self-administered, self-assessed usual health status assessment, which was obtained from the HUI developers and covers both Mark II and Mark III health status classification systems.35 The Mark III classification system is now recommended by the developers because of its broad applicability in both clinical and general population health studies, improvements in a number of definitions, and an increased orthogonality of its attributes for structural independence.36 The HUI Mark III health status classification system covers 8 attributes: cognition, vision, hearing, speech, ambulation, dexterity, emotion, and pain. Function within each attribute is graded on a 5- or 6-point scale that corresponds to level of severity, ranging from normal function to severe impairment. Responses to the HUI Mark III can be converted into a utility score by reference to a utility scoring algorithm that can be attached to each permutation of responses.37 The utility scoring algorithm can be summarized as $u^* = 1.371 + \sum b_i$, where $u^*$ is the utility of a chronic health state on the 0–1 scale and $b_i$ are coefficients provided by the HUI developers for the appropriate attribute and level.37 For development of the utility scoring algorithm, a random sample of 504 adults who were general population and living in the city of Hamilton, Canada, had previously been asked to value selected health states using both a visual analog scaling technique and a standard gamble instrument.38 Additional details on the utility algorithm for the HUI Mark III are reported elsewhere.36

Ethical Approval

The ELGA study was approved by the Oxford Multi-Centre Ethics Committee and by local ethics committees in the former Northern Region and in Scotland.

Statistical Analyses

In this analysis, we used the self-reports of the teenagers only. Comparisons were made between the ELGA teenagers in mainstream schools and their term-born control subjects. Differences in baseline characteristics between the ELGA children and their control subjects were tested using the Pearson $\chi^2$ test. For each of the 8 attributes of the HUI Mark III, we compared the proportion of children with (1) any level of functional limitation and (2) severe functional impairment using Fisher's exact test for equality of proportions. Severe functional impairment for each of the 8 attributes of the HUI Mark III had previously been defined by the HUI developers to facilitate comparisons among patient groups and over time.35 Severe functional impairment was defined as levels 5 and 6 in the cognition, vision, and hearing attributes; level 5 in the speech attribute; levels 4, 5, and 6 in the ambulation and dexterity attributes; and levels 4 and 5 in the emotion and pain attributes. Descriptive statistics were calculated for the derived utility scores. Differences in utility scores were tested using 2-sample $t$ tests for unequal variance. Because many comparisons were made, we used a conservative $P$ value threshold of <.01. Statistical analyses were conducted using Stata 8.0 (Stata Corp, College Station, TX).

RESULTS

Of the 535 children in the original ELGA cohorts, 218 were alive at 15 to 16 years of age and approached to participate in the study. This rate of mortality was not uncommon in the presurfactant era of neonatology. The postal questionnaires were returned by 175 (80.3%) ELGA children, 147 of whom were in mainstream schools at the time of the study and 28 of whom were in special needs schools. In the United Kingdom, the majority of children with special needs are educated in mainstream schools. However, there are a few children whose needs are so complex that they require education outside mainstream education in special schools. A complete HUI Mark III record was available for 140 of the 147 ELGA children in mainstream schools, as well as for all 108 control subjects approached. These children formed the basis of our analyses. Nonresponders in the ELGA group were comparable to responders in terms of birth weight and gestational age but were of lower social status (data available on request). The demographic and
socioeconomic characteristics of the 140 participating ELGA children in mainstream schools and the 108 control subjects were broadly similar (Table 1).

Comparisons of the frequency and the proportion of any functional impairment between the mainstream ELGA teenagers and control subjects are shown in Table 2 for each of the 8 attributes of the HUI Mark III. In 7 of the 8 attributes (vision, hearing, speech, emotion, pain, ambulation, and dexterity), there were no statistically significant differences in any functional impairment between the ELGA children in mainstream schools and their control subjects. However, the ELGA children in mainstream schools did report a higher level of functional impairment in the cognition attribute (40.7% vs 25.0%; \( P = .010 \)). When the analyses were restricted to frequencies of severe functional impairment, there were no statistically significant differences between the comparison groups across all 8 attributes of the HUI Mark III (Table 3).

Table 4 gives a description of the overall HUI Mark III utility scores for the comparison groups. The mean utility score for the ELGA children in mainstream schools was 0.86, compared with 0.89 for the control group, a mean difference in utility score of 0.03 that was not statistically significant (\( P = .123 \)). Although there was no difference in the median utility score between the 2 groups (0.93), there was a broader range of utility scores for the ELGA group (0.07–1.0 vs 0.45–1.0 for the control group). A total of 3 (2.1%) ELGA children had a utility score of between 0 and 0.25, and 6 (4.3%) had a utility score of between 0.26 and 0.50. The respective numbers for the control group were 0 and 1 (0.9%; Table 5).

### TABLE 1 Sociodemographic Characteristics of Mainstream ELGA and Control Teenagers

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mainstream ELGA Teens (( N = 140 ), n (%))</th>
<th>Control Teens (( N = 108 ), n (%))</th>
<th>( P^a )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td>0.286</td>
</tr>
<tr>
<td>Male</td>
<td>64 (45.7)</td>
<td>42 (38.9)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>75 (53.6)</td>
<td>65 (60.2)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>1 (0.7)</td>
<td>1 (0.9)</td>
<td></td>
</tr>
<tr>
<td>Social status (Registrar General Classification( ^b ))</td>
<td></td>
<td></td>
<td>0.535</td>
</tr>
<tr>
<td>1 and 2</td>
<td>37 (26.4)</td>
<td>37 (34.3)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>53 (37.9)</td>
<td>44 (40.7)</td>
<td></td>
</tr>
<tr>
<td>4 and 5</td>
<td>18 (12.9)</td>
<td>11 (10.2)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>32 (22.9)</td>
<td>16 (14.8)</td>
<td></td>
</tr>
<tr>
<td>Parental qualification</td>
<td></td>
<td></td>
<td>0.468</td>
</tr>
<tr>
<td>No qualifications (left school at age 16)</td>
<td>47 (33.6)</td>
<td>30 (27.8)</td>
<td></td>
</tr>
<tr>
<td>High School (&quot;O&quot; or &quot;A&quot; levels)</td>
<td>64 (45.7)</td>
<td>46 (42.6)</td>
<td></td>
</tr>
<tr>
<td>University degree</td>
<td>24 (17.1)</td>
<td>24 (22.2)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>5 (3.6)</td>
<td>8 (7.4)</td>
<td></td>
</tr>
</tbody>
</table>

\( ^a \) Calculated using Pearson \( \chi^2 \) test, excluding missing values.

### TABLE 2 Frequency of Any Level of Functional Impairment Within Each HUI Attribute for Mainstream ELGA and Control Teenagers According to Teenager Self-report

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mainstream ELGA Teens (( N = 140 ), n (%))</th>
<th>Control Teens (( N = 108 ), n (%))</th>
<th>( P^a )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vision</td>
<td>38 (27.1)</td>
<td>25 (23.1)</td>
<td>.557</td>
</tr>
<tr>
<td>Hearing</td>
<td>1 (0.7)</td>
<td>0</td>
<td>.999</td>
</tr>
<tr>
<td>Speech</td>
<td>19 (13.6)</td>
<td>13 (12.0)</td>
<td>.849</td>
</tr>
<tr>
<td>Emotion</td>
<td>40 (28.6)</td>
<td>37 (34.3)</td>
<td>.406</td>
</tr>
<tr>
<td>Pain</td>
<td>31 (22.1)</td>
<td>39 (36.1)</td>
<td>.022</td>
</tr>
<tr>
<td>Ambulation</td>
<td>5 (3.6)</td>
<td>1 (0.9)</td>
<td>.237</td>
</tr>
<tr>
<td>Dexterity</td>
<td>6 (4.3)</td>
<td>0</td>
<td>.037</td>
</tr>
<tr>
<td>Cognition</td>
<td>57 (40.7)</td>
<td>27 (25.0)</td>
<td>.010</td>
</tr>
</tbody>
</table>

\( ^a \) Calculated using Fisher's exact test.

### TABLE 3 Frequency of Severe Functional Impairment Within Each HUI Attribute for Mainstream ELGA and Control Teenagers According to Teenager Self-report

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mainstream ELGA Teens (( N = 140 ), n (%))</th>
<th>Control Teens (( N = 108 ), n (%))</th>
<th>( P^a )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vision</td>
<td>3 (2.1)</td>
<td>0</td>
<td>.260</td>
</tr>
<tr>
<td>Hearing</td>
<td>0</td>
<td>0</td>
<td>.999</td>
</tr>
<tr>
<td>Speech</td>
<td>1 (0.7)</td>
<td>0</td>
<td>.999</td>
</tr>
<tr>
<td>Emotion</td>
<td>1 (0.7)</td>
<td>1 (0.9)</td>
<td>.999</td>
</tr>
<tr>
<td>Pain</td>
<td>1 (0.7)</td>
<td>2 (1.9)</td>
<td>.582</td>
</tr>
<tr>
<td>Ambulation</td>
<td>3 (2.1)</td>
<td>0</td>
<td>.260</td>
</tr>
<tr>
<td>Dexterity</td>
<td>3 (2.1)</td>
<td>0</td>
<td>.260</td>
</tr>
<tr>
<td>Cognition</td>
<td>15 (10.7)</td>
<td>9 (8.3)</td>
<td>.666</td>
</tr>
</tbody>
</table>

\( ^a \) Calculated using Fisher's exact test.

**DISCUSSION**

Previous studies of the long-term health outcomes of children who were born preterm have tended to focus on relatively narrow biomedical measures of morbidity. Recently, investigators from a number of disciplines, including anthropology, economics, sociology, and psychology, have recognized the importance of measuring the impact of preterm birth across multiple domains. A particular approach to measuring multiple health outcomes is the multiattribute utility measure, which generates not only scores across disparate attributes of health but also an overall score on a scale from 0 to 1 that reflects population or patient preferences for the overall health state that is being measured. This study uses the HUI Mark III to describe the self-reported health status and health-related quality of life of British teenagers who are in mainstream schooling and were born at ELGA and a comparison group of teenagers who were born at term. It reveals no statistically significant differences in any functional impairment in 7 of the 8 attributes considered between the preterm adolescents in mainstream schools and their control subjects. In addition, there was no significant difference in the overall utility score between the 2 groups, although the preterm adolescents did display a broader range of utility scores.
The results of this study are broadly in line with what has been reported elsewhere in the published literature. However, it is worth noting that most of these studies are based on preterm samples with higher gestational age than this sample. A recently conducted comprehensive review of studies of health-related quality of life of preterm children concluded that the majority of these children do not rate their health-related quality of life as significantly different from that of term-born control subjects despite objective evidence that they have poorer health on average. Furthermore, this finding has been replicated, in large part, by a number of empirical studies. Danish researchers assessed the health-related quality of life of 85 young adults (18–20 years of age), born in 1971–1974 with birth weights <1500 g, and made comparisons with 85 young adults who were born at >2500 g during the same period. Health-related quality of life was assessed by telephone interview using an instrument that covered elementary biological needs, warm human relationships, meaningful occupation, and diverse and exciting experiences. Young adults who had birth weights of <1500 g and reported no impairments did not differ from control subjects on scores of health-related quality of life, but those who reported physical or mental impairment had scores of health-related quality of life that were significantly lower than those in the control group. These researchers repeated their study on a cohort born in 1980–1982. The results were essentially identical, although those with birth weights of <1500 g reported significant impairment on objective as opposed to subjective assessment of health-related quality of life. A Swedish study compared the health-related quality of life of 39 young adults (age 19) who were born before 35 weeks’ gestation and 23 term-born control subjects. Self-rated health-related quality of life was assessed using the visual analog scaling technique. The investigators found no significant differences between the 2 groups.

In contrast to the findings above, a series of studies by Saigal et al of ELBW children did find significant differences in health status and health-related quality of life compared with control subjects. In an early study by these investigators, the health-related quality of life of ELBW children and a reference group of children at 8 years of age was retrospectively classified using the HUI Mark II classification on the basis of assessments that were provided by health professionals. The utility scores for each child were estimated indirectly using a formula that was derived from preference measurements that were obtained from parents in the general population. Mean utility scores were lower for the ELBW children than for the reference group. A later study by Saigal et al compared the health status and health-related quality of life of 141 children who born weighing <1000 g and 145 normal birth weight control subjects. Children completed the HUI Mark II classification between ages 12 and 16. The children who born weighing <1000 g were significantly more limited in cognition, sensation, self-care, and pain compared with the control subjects. Furthermore, they had a significantly lower mean utility score. However, the vast majority of the ELBW group viewed their health-related quality of life as satisfactory, and it was difficult to distinguish their scores from those of the term-born control subjects. It should be noted that there were differences between the methods used by the latter study and those used by our study. Notably, the research instruments in the study by Saigal et al were interviewer administered, and the utility scores that were attached to the HUI Mark II responses were obtained directly from the children themselves using the visual analog and standard gamble techniques. In contrast, the children in our study completed postal questionnaires and the utility scores that were attached to the HUI Mark III responses were derived from a general population of Canadian adults. Recent research suggests that our approach of indirectly estimating utility scores by attaching population-derived utility scores to HUI health states may be a poor substitute for directly measured utility scores. Nevertheless, there is no evidence to suggest that this systematically biases group differences in utility scores.

### Table 4

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean (SD)</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean Difference</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELGA teens</td>
<td>140</td>
<td>0.86 (0.19)</td>
<td>0.93</td>
<td>0.07</td>
<td>1.0</td>
<td>0.03</td>
<td>.123</td>
</tr>
<tr>
<td>Control subjects</td>
<td>108</td>
<td>0.89 (0.12)</td>
<td>0.93</td>
<td>0.45</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Calculated using 2-sample t test for unequal variance.

### Table 5

<table>
<thead>
<tr>
<th>Range of Utility Scores</th>
<th>Mainstream ELGA Teenagers (N = 140, n (%))</th>
<th>Control Teenagers (N = 108, n (%))</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>228</td>
</tr>
<tr>
<td>0.00–0.25</td>
<td>3 (2.1)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0.26–0.50</td>
<td>6 (4.3)</td>
<td>1 (0.9)</td>
<td></td>
</tr>
<tr>
<td>0.51–0.75</td>
<td>18 (12.9)</td>
<td>11 (10.2)</td>
<td></td>
</tr>
<tr>
<td>0.76–1.00</td>
<td>113 (80.7)</td>
<td>96 (88.9)</td>
<td></td>
</tr>
</tbody>
</table>

*Calculated using Pearson χ² test.
There seem to be 3 main explanations as to why many studies of health status and health-related quality of life, including this one with an extremely preterm cohort, fail to find major differences between preterm-born teenagers and term-born control subjects. First, it is possible that the measures of health status and health-related quality of life that are applied in these studies are poor. The limited evidence that is available presents a mixed picture of the psychometric properties of the alternative measures of health status and health-related quality of life applied in a childhood context. It is evident that additional research is required to establish the practicality, reliability, and validity of these measures when applied to pediatric populations. In particular, it may be that the reading level that is required for the HUI and other measures is somewhat high for pediatric samples in which a number of children may have mild learning difficulties. A second explanation for the non-significant differences between our study groups is that preterm-born teenagers who attend mainstream schooling are likely to be healthier and have less disability than those who require special schooling. In particular, the ability to cope in a mainstream school usually precludes those with severe cognitive or behavioral problems. Therefore, it may not be surprising that this group of preterm children are little different from their peers. However, in a previous report on these teens, parents of ELGA teenagers in mainstream school reported a higher incidence of problems than did parents of control subjects in physical functioning, mental health, and family life. Furthermore, their teachers rated their ability lower than that of the control group. A final consideration when interpreting our results is that it may be possible that teenagers’ perceptions of their health status and health-related quality of life may be different from that of their families, caregivers, or health professionals. In a companion article that examined emotional and behavioral problems in the same population, group differences in conduct and emotional problems were less pronounced when teens self-reported on their own well-being, compared with when these reports came from teachers and parents, suggesting that young people might rate their problems differently from adults who know them well. However, youth perceptions would not seem to account for the lack of group differences in these data, because the HUI responses that were reported by the ELGA preterm and control group parents were almost identical (data available from authors).

The only significant difference in health status between the ELGA teenagers and control subjects was on the cognition attribute. This attribute is tested using 2 questions: 1 on memory and 1 on thinking and problem-solving. Previous studies have established that very low birth weight infants are more likely to develop visual perceptual and visual-motor impairments, delay in some language functions, and working memory deficit, and, at school age, they may have learning problems and attention deficit, although few have focused on specific abilities of the cognitive spectrum during late childhood. The mechanisms by which prematurity might cause cognitive deficits in adolescence require additional research.

There are a number of strengths to the study reported in this article. First, it is based on an extremely preterm population-based cohort that was drawn from defined geographic areas rather than a clinic-based population; consequently, selection biases are unlikely to represent a major problem. Second, children were recruited from 3 regions of the United Kingdom that reflect its socioeconomic and ethnic diversity; therefore, the study is likely to have high external validity. Third, the study used a validated and reliable measure of health status and health-related quality of life, HUI. Fourth, the analysis used school-age control subjects that were specifically recruited for this study rather than data from siblings, which are prone to biases as a result of continuously changing developmental situations, or comparisons with British population norms, for which limited data are available.

The study does have a number of caveats, which should be borne in mind by readers. First, it was not possible to identify an appropriate comparison group for the 28 index children in special needs schools, because classes in special needs schools are organized according to ability to participate in particular activities rather than age. Therefore, these children were not included in this analysis. Separate comparisons between the entire ELGA cohort, including the children who were attending special needs schools, and the 108 term-born control subjects revealed that the ELGA children as a totality did report a statistically significant higher level of functional impairment in 4 of the 8 HUI Mark III attributes (cognition, hearing, ambulation, and dexterity). Furthermore, the mean utility score was significantly lower for the entire ELGA cohort when compared with the term-born control subjects (0.82 vs 0.89; P = .003). Nevertheless, given the absence of control subjects for the 28 index children in special needs schools, we believe that the appropriate comparisons are those that we report in the main section of the article. Second, although the response rates to our questionnaires were relatively high, there is a concern that the outcome of the children who were not seen or assessed differs from those who were seen. Additional analyses revealed that the nonresponders to our study included a higher proportion of children without previous assessments and a higher proportion of children of lower social status. This obviously needs to be borne in mind when interpreting the group findings. Third, although the HUI is the most widely used of the multiattribute utility measures within the childhood context, the underlying preference weights that were provided by the developers were derived from a survey of Canadian adults. Additional research that
elucidates the underlying valuations of preterm children with the necessary cognitive capacities for the health states that they experience during adolescence and later into adulthood is required. Fourth, although it is possible that bias was introduced though the year tutor’s choice of control subjects, this is unlikely, because they were requested to choose the children who were closest in date of birth to the ELGA child.

CONCLUSIONS
Follow-up studies of very preterm infants into adolescence and adulthood have benefited tremendously from the examination of a broader set of outcomes than simply morbidity and mortality. In particular, measures of health-related quality of life and health utilities have enabled the translation of neurologic and other sequelae into outcomes that are arguably more meaningful to patients, health service providers, and policy makers. These comprehensive accounts of individual patient experience provide a more detailed account on which to base service provision and to evaluate outcome of interventions. They also permit comparison with other chronic conditions of childhood, which can help with prioritization and funding of services. Their use in follow-up studies, however, is not yet routine, and more methodologic work is required, particularly with regard to eliciting underlying preferences for the health states that are experienced from the children themselves and in understanding the meaning of young peoples’ self-perceptions of their health.

ACKNOWLEDGMENTS
This study was supported by core funding for the National Perinatal Epidemiology Unit from the Department of Health in England.

We acknowledge the contribution of the participants in this project, the researchers who invested a great deal of time and effort in setting up this study and following up the children, and the ELGA principal investigators for giving permission for this analysis. The original ELGA group of principal investigators consisted of Frances Gardner, Edmund Hey, Ann Johnson, Lesley Mutch, Unni Wariyar, and Patricia Yudkin. The ELGA steering group consisted of the investigators plus Sarah Arkle, Ursula Bowler (project coordinator), Christine Hockley, Michael Jones, Barbara Maughan, and Anne Stewart.

REFERENCES


46. McCormick MC. The outcomes of very low birth weight infants: are we asking the right questions? *Pediatrics* 1997;99:869–876

OBJECTIVES. Our goal was to compare BMI and waist circumference with dual-energy radiograph absorptiometry–based measures of adiposity and to describe the pattern and interrelations of these surrogate and direct adiposity measures in prepubertal and pubertal rural Chinese children.

METHODS. This was a cross-sectional study of 2493 children aged 6 to 18 years from a population-based cohort of twin pairs. Dual-energy radiograph absorptiometry–based measurements included total body fat, percentage of body fat, trunk fat, and percentage of trunk fat. Age- and gender-specific patterns and interrelationships among BMI, waist circumference, and dual-energy radiograph absorptiometry–based measurements were described by using smoothing plots and age- and gender-specific correlation analyses.

RESULTS. In girls, BMI, waist circumference, total body fat, percentage of body fat, trunk fat, and percentage of trunk fat all increased linearly with age. In boys, BMI and waist circumference increased linearly with age, but total body fat, percentage of body fat, and trunk fat did not increase significantly with age. In both genders, percentage of trunk fat reached a nadir around 12 years of age and then increased with age. Before puberty (6–11 years), BMI and waist circumference were correlated well with total body fat, percentage of body fat, and trunk fat in both genders. During puberty (12–18 years), the correlations between BMI and each of the dual-energy radiograph absorptiometry–based measurements were higher in girls than in boys. Similar trends were found in the correlations between waist circumference and each of the dual-energy radiograph absorptiometry–based measurements.

CONCLUSIONS. In this relatively lean rural Chinese population, BMI and waist circumference were highly correlated with each other and were good surrogates of total body fat, trunk fat, and percentage of body fat in prepubertal children of both genders and in pubertal girls. However, both BMI and waist circumference overestimated total and trunk fat, especially percentage of body fat in pubertal boys.
CHILDHOOD AND ADOLESCENT obesity is increasing worldwide. From 1992 to 2002, the prevalence of overweight and obesity in Chinese people aged 0 to 6 and 7 to 17 years increased by 31.7% and 17.9%, respectively.\textsuperscript{1,2} Therefore, obesity and its associated cardiovascular disease risk factors are emerging as important public health issues for children and adolescents.\textsuperscript{3} Exposure to obesity early in life may induce arterial changes contributing to the development of atherosclerosis in adulthood.\textsuperscript{4} Although there is a lower prevalence of obesity in rural areas of China, the rate of increase in obesity in these areas is outpacing that in urban areas.\textsuperscript{1,2} Rural residents constitute a large segment of the world’s total population, particularly in developing countries. In China, 85% of the population lives in rural, agricultural regions.\textsuperscript{5} Because of these trends, the prevalence of overweight and obesity in China will likely increase rapidly, as will the illnesses associated with obesity. The public health costs associated with these changes will be immense.

This study simultaneously evaluated surrogate adiposity measurements (BMI, waist circumference [WC]) and direct adiposity measures derived from dual-energy radiograph absorptiometry (DEXA). BMI, based on height and weight measurements, is routinely obtained in clinical settings and used to assess overweight and obesity in children and adults.\textsuperscript{6} WC is also easily obtained in medical practice as a sign of central obesity and used as a characteristic of the metabolic syndrome. DEXA is one of the best available measures of body fatness.\textsuperscript{7–10} In this study the body fat measures from DEXA include total body fat (TBF), percentage of body fat (%BF), trunk fat (TF), and percentage of trunk fat (%TF).

This study has 2 aims. First, it compares standard field measures of BMI and WC to laboratory measures of TBF, %BF, TF, and %TF in prepubertal and pubertal rural Chinese children. BMI and WC are often used to identify individuals and populations with increased adiposity to target for public health interventions. Our study will help determine whether these field measures can serve as good surrogates for laboratory measures of adiposity. Second, the study describes the pattern and interrelations of these surrogate and direct adiposity measures in a large sample of rural Chinese children. In particular, this study assesses the differential patterns and interrelations between prepubertal and pubertal children and between boys and girls, given the known differences in growth, development, and body composition by gender during the pubertal period.

METHODS

Study Population and Recruitment
This study uses a subset of subjects from a large cohort of twin pairs in Anqing, China, recruited from 1998 to 2000, with the goal to study environmental and genetic determinants of complex human diseases, including obesity and metabolic syndrome. Spanning 80 km along the north bank of the Yangtze River, the area of Anqing has 3 urban areas and 8 rural counties covering 15 000 km.\textsuperscript{2} Medical care in each county of Anqing is administered through a 3-tier (county, township, and village) service network. Twins were identified through a multistage process. First, investigators from Anhui Medical University and the Anqing Hospitals/Research Institutes held a 3-day workshop in each township to train local doctors to participate in subject recruitment. The first day was used to explain the purpose, scope, and procedures of the study. The definition of a twin was introduced, and several examples were presented. Local doctors were requested to go back to their own villages to prepare a list of all of the twins in their practice area. Epidemiologists from Anhui Medical University checked all of the twin lists with the township/village doctors. Twins were chosen on the basis of the following criteria: (1) age of 6 to 60 years, (2) both twins were available for the survey, and (3) both twins (or parents/guardians of children) agreed and consented to participate in the survey. Eligible twins were invited to a central office to complete a questionnaire interview, blood drawing, and physical examination, including anthropometric measurements and DEXA scan. All of the study protocols were approved by the institutional review boards of Children’s Memorial Hospital and Anhui Medical University.

Anthropometric Measurements
A detailed description of field-data collection has been described elsewhere.\textsuperscript{11,12} In brief, height and weight were measured using standard protocols, without shoes or outerwear. Height was measured to the nearest 0.1 cm on a portable stadiometer. Weight was measured to the nearest 0.1 kg with the subjects standing motionless on a scale, and BMI was calculated as weight (kilograms)/height squared (meters squared). A WC measurement was taken at the level of the umbilicus to the nearest millimeter. Each anthropometric measure was taken 3 times and the mean used in all of the analyses.

DEXA
DEXA measures the exponential attenuation of photons emitted at 2 energy levels that are absorbed by various body tissues. This allows for accurate measurements of fat, fat-free, and bone substances.\textsuperscript{13} A standard whole-body DEXA scan includes total body and 3 regional fat measures: trunk (chest, abdomen, and pelvis), arms, and legs. DEXA measures of body fat have been validated against other estimates, including underwater weighing,\textsuperscript{8} skinfold measures, bioelectrical impedance analysis, and deuterium oxide dilution.\textsuperscript{10} A standard software calculation\textsuperscript{13} was used to calculate TBF measured with a
Lunar DPXL instrument (Madison, WI) that was set up in Anqing. %BF is calculated as TBF divided by body weight. %TF is calculated as trunk fat divided by TBF and is a measure of central fat distribution relative to TBF.

Statistical Methods

Our analysis compares BMI and WC to DEXA-based measures of adiposity and describes the interrelationships of body mass and fat patterning in prepubertal and pubertal children. Twins were treated as individuals from the general population in our analyses. All of the statistical analyses were conducted using the SAS statistical package (SAS Institute, Inc, Cary, NC). Study participants are grouped into 2 age groups: 6 to 11 years (approximating prepubertal) and 12 to 18 years (approximating pubertal). The age of menarche in girls in this population is 13.9 years of age. The onset of puberty is, in general, 2 years before menarche; thus, 12 years of age was chosen as the cutoff in our population.

All of the analyses are gender- and age-group specific and were conducted in several steps. First, the distributions of BMI, WC, TBF, %BF, TF, and %TF in both genders were examined. The relationships between age and body-fat measures are described using smoothing plots used locally weighted regression. Furthermore, the nonparametric method for smoothing with SAS procedure locally weighted regression as a smoothing plots used locally weighted regression.

The relationships among BMI, WC, and body-fat measures are described using smoothing equations. The original twin cohort enrolled a total of 3412 children between 6 and 18 years of age. In this report, 919 children were excluded because of missing data for DEXA measures \((n = 871)\), BMI and WC measures \((n = 35)\), reported history of smoking \((n = 9)\), and drinking alcohol \((n = 4)\). Thus, the final analyses included 2493 children. The age and gender distribution and BMI are comparable between the 2493 children included and 919 children excluded from the analyses (data not shown). Table 1 displays gender and age group-specific means and SDs for all of the variables used in the analyses, including age, height, weight, BMI, WC, and DEXA measures of adiposity (TBF, %BF, TF, and %TF).

### Results

#### Epidemiologic Characteristics

The original twin cohort enrolled a total of 3412 children between 6 and 18 years of age. In this report, 919 children were excluded because of missing data for DEXA measures \((n = 871)\), BMI and WC measures \((n = 35)\), reported history of smoking \((n = 9)\), and drinking alcohol \((n = 4)\). Thus, the final analyses included 2493 children. The age and gender distribution and BMI are comparable between the 2493 children included and 919 children excluded from the analyses (data not shown). Table 1 displays gender and age group-specific means and SDs for all of the variables used in the analyses, including age, height, weight, BMI, WC, and DEXA measures of adiposity (TBF, %BF, TF, and %TF).

#### Relationship Between Age and Body-Fat Measures

Figure 1 displays locally weighted regression smoothing plots that show the associations between body-fat measures and age, stratified by gender. BMI and WC increase linearly throughout the age range in both genders but accelerate during puberty. TBF increases in both genders throughout the age range but diverges during puberty when TBF increases more rapidly in girls. Of note, %BF increases in boys from 6 to 12 years of age but decreases during puberty; in girls, %BF increases slowly in prepuberty and accelerates during puberty, which shows remarkable divergence by gender during puberty. Trunk fat increases linearly in boys throughout the age range but accelerates sharply at puberty in girls. In both genders, the %TF has a J-shaped curve, decreasing during prepuberty (to age 10 in girls and to age 11 in boys) and increases sharply in both genders during puberty.

### TABLE 1

<table>
<thead>
<tr>
<th>Variables</th>
<th>6–11 y</th>
<th>12–18 y</th>
<th>P&lt;sup&gt;a&lt;/sup&gt;</th>
<th>6–11 y</th>
<th>12–18 y</th>
<th>P&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>9.37 ± 1.37</td>
<td>9.56 ± 1.40</td>
<td>.0073</td>
<td>14.30 ± 1.72</td>
<td>14.65 ± 2.09</td>
<td>.0040</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>23.40 ± 4.22</td>
<td>23.21 ± 4.55</td>
<td>.3881</td>
<td>39.74 ± 9.75</td>
<td>39.64 ± 8.34</td>
<td>.8527</td>
</tr>
<tr>
<td>Height, m</td>
<td>1.04 ± 0.08</td>
<td>1.25 ± 0.09</td>
<td>.5788</td>
<td>1.50 ± 0.11</td>
<td>1.47 ± 0.08</td>
<td>.0004</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>15.03 ± 1.43</td>
<td>14.81 ± 1.47</td>
<td>.0043</td>
<td>17.51 ± 2.43</td>
<td>18.08 ± 2.63</td>
<td>.0003</td>
</tr>
<tr>
<td>WC, cm</td>
<td>52.74 ± 4.08</td>
<td>51.57 ± 4.64</td>
<td>&lt;.0001</td>
<td>60.09 ± 5.69</td>
<td>60.60 ± 6.05</td>
<td>.0546</td>
</tr>
<tr>
<td>TBF, kg</td>
<td>2.04 ± 1.07</td>
<td>2.67 ± 1.28</td>
<td>&lt;.0001</td>
<td>3.84 ± 2.01</td>
<td>7.85 ± 4.18</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>%BF, %</td>
<td>8.47 ± 3.08</td>
<td>11.18 ± 3.95</td>
<td>&lt;.0001</td>
<td>9.63 ± 3.80</td>
<td>18.72 ± 6.61</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>TF, kg</td>
<td>0.73 ± 0.41</td>
<td>0.94 ± 0.54</td>
<td>&lt;.0001</td>
<td>1.56 ± 0.99</td>
<td>3.40 ± 2.18</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>%TF, %</td>
<td>36.24 ± 4.93</td>
<td>34.97 ± 4.84</td>
<td>&lt;.0001</td>
<td>39.97 ± 5.97</td>
<td>41.00 ± 6.21</td>
<td>.0074</td>
</tr>
</tbody>
</table>

Results are mean ± SD.

<sup>a</sup> P is a 2-sided comparison of boys’ and girls’ means by t test.
Relationship of BMI with WC, %BF, and %TF

Figure 2 depicts gender and age group-specific relationships of BMI with WC, %BF, and %TF. There were few subjects with low BMI (BMI <11 kg/m² in ages 6–11 years and BMI <13 kg/m² in ages 12–18 years) or high BMI (BMI ≥20 kg/m² in ages 6–11 years and ≥23 kg/m² in ages 12–18 years) in this population. Thus, in the analysis of children aged 6 to 11 years, subjects with BMI ≤11 kg/m² were analyzed as having BMI of 11 kg/m². Likewise, subjects with BMI ≥20 kg/m² were analyzed
as having BMI of 20 kg/m². In the analysis of ages 12 to 18, subjects with BMI ≤13 kg/m² were analyzed as having BMI of 13 kg/m², and subjects with BMI ≥23 kg/m² were analyzed as having BMI of 23 kg/m².

There is a linear increase in WC with increasing BMI in both age groups and genders. In children aged 6 to 11 years, there is a linear increase in %BF with increasing BMI in both boys and girls. However, there are clear gender differences in the relationship of BMI with %BF in children aged 12 to 18 years. For girls aged 12 to 18 years, there is a linear increase in BMI and %BF. However, in boys there is no apparent increase in %BF with increased BMI until the BMI is ≥20 kg/m².

%TF in girls aged 6 to 11 years is more or less constant.
until a BMI ≤ 16 kg/m², at which point it increases sharply. In boys, %TF decreases with increasing BMI until a BMI of 16 kg/m², at which point it increases. In both genders aged 12 to 18 years, there is a linear relationship between BMI (only for BMI > 16 kg/m² in boys) and %TF. Of interest, for both age groups, the %TF of boys is higher than that of girls if BMI is < 16 kg/m² but lower than that of girls if BMI is > 16 kg/m². The relationships between WC and %BF and %TF were similar to those of BMI in both age groups and genders (figures not shown).

Correlation Between Body-Fat Measures
Table 2 demonstrates that in children aged 6 to 11 years, BMI and WC are similarly correlated with the corresponding body fat as measured by DEXA in both genders. The correlation coefficients between BMI and body fat from DEXA (TBF, %BF, and TF) are 0.57, 0.43, and 0.55 in boys and 0.57, 0.40, and 0.54 in girls, respectively. %TF has no correlation with BMI in boys and has mild correlation with BMI in girls 6 to 11 years of age. There is a gender difference in correlation among body-fat measures in children aged 12 to 18 years (P < .05; data not shown). The correlation coefficients between BMI and each of TBF, %BF, TF, and %TF are 0.57, 0.19, 0.63, and 0.49 in boys and 0.88, 0.80, 0.88, and 0.69 in girls, respectively. A similar trend is found in the correlation between WC and these direct body-fat measures by DEXA.

DISCUSSION
This report contributes new information on adiposity and body composition in children and adolescents in several ways. It is one of the first studies to describe gender-specific patterns and interrelationships of body-fat measures in a large rural Chinese population using DEXA-derived measures from childhood through adolescence. It examines the use of BMI and WC, the most commonly used surrogate measures of adiposity, to assess adiposity from prepuberty through puberty in this population. Finally, the results highlight gender and pubertal stage as important determinants in adiposity measurements and their interrelationships.

Our data showed important differences in measures of body fat between boys and girls with the onset of puberty. In girls, BMI, WC, TBF, %BF, and TF all increased linearly with age with a more pronounced increase observed at the onset of puberty (12 years of age). In boys, TBF and TF increased slightly with age. %BF also increased slightly with age during prepuberty, but at 12 years of age and onward, %BF began to decrease with age despite the fact that BMI continued to increase.

During prepuberty, boys had higher BMI, WC, and %TF than girls. After puberty the BMI and WC of boys are only slightly lower than those of girls. In contrast, DEXA measures of body fat in girls (with the exception of %TF) are higher than in boys throughout the entire age range and become more pronounced with the onset of puberty. In adolescence, boys have an increase in muscle mass and central adipose tissue because of testosterone secretion, whereas girls have increased body-fat mass, primarily peripheral adipose tissue, because of estradiol. The increased adiposity seen in girls that accelerates during puberty is consistent with these changes. These gender differences in fat patterning with puberty result in boys and girls with very different body composition. Thus, it is not surprising that boys and girls in our population have very different levels of adiposity as measured by DEXA when their BMI and WC are very similar.

There was a similar level of central fat distribution in boys and girls from 6 to 18 years of age despite the fact that girls had higher TBF than boys during puberty. These changes are likely because of sexual dimorphism in body-fat distribution mediated by estradiol and testosterone. In boys, there is an accumulation of both subcutaneous and intraabdominal fat in the upper body in an apple-shaped distribution. In contrast, girls have a gluteal accumulation of subcutaneous fat in a pear-shaped distribution. Some studies show that these changes start in adolescents, whereas others report sexual dimorphism in children as young as 5 years of age. The cause of prepubertal sexual dimorphism in fat patterning is not known. Garnett et al. investigated the role of insulin-like growth factor 1, dehydroepiandrosterone sulfate, estradiol, testosterone, and leptin on DEXA-measured %BF and percentage of abdominal fat and found that the hormones examined explained <20% of gender differences in fat patterning. Hormonal regulators of adipose tissue need to be further studied to account for gender differences in body-fat patterning. Ethnic differences may also affect fat distribution during puberty. One study found progression of sexual dimorphism in fat patterning with increasing pubertal

| Table 2 The Pearson Correlation Coefficients Among BMI, WC, and Body Fat According to Age Groups and Genders |
|---------------------------------------------------|-----------------|-----------------|
| Body-Fat Measures From DEXA | 6–11 y BMI WC | 12–18 y BMI WC |
| TBF | | |
| Boys | 0.57 | 0.49 | 0.57 | 0.60 |
| Girls | 0.57 | 0.42 | 0.88 | 0.76 |
| %BF | | |
| Boys | 0.43 | 0.36 | 0.19 | 0.24 |
| Girls | 0.40 | 0.27 | 0.80 | 0.70 |
| TF | | |
| Boys | 0.55 | 0.52 | 0.63 | 0.66 |
| Girls | 0.54 | 0.43 | 0.88 | 0.78 |
| %TF | | |
| Boys | -0.04 | 0.10 | 0.49 | 0.51 |
| Girls | 0.17 | 0.20 | 0.69 | 0.69 |

*P = .0029, whereas all other P values are <.01.*
maturation in Asian children.24 A study on German children found that the developmental pattern of fat accumulation and distribution during adolescence is highly dynamic and gender specific.27 It is not surprising that studies have differing results, because age, gender, race, study methods, and the relative leanness of a population may play a role.

This study examined the interrelationships of body-fat measures, specifically whether BMI and WC can be used as surrogates of DEXA-based measurements of body fat. Our data demonstrate that, during puberty, there were gender differences in the relationship of BMI and WC to DEXA-based measures of body fat. Among girls, BMI and WC correlate well with DEXA measure of adiposity, and the correlations increase considerably from preadolescents to adolescence. The correlations likely improved because of the higher levels of adiposity in puberty. In contrast, BMI and WC did not accurately reflect adiposity in boys aged ≥12 years in this rural Chinese population. This observation could be explained by differential body composition between boys and girls seen during puberty such that, for a given BMI, boys are likely to have more lean muscle mass than girls, resulting in BMI overestimating adiposity in adolescent boys. However, in studies of American and Italian children and adolescents, BMI was highly correlated with %BF in both genders.28 This Chinese cohort is leaner than the others, and because the correlation between BMI and %BF is stronger in heavier children, the leanness of the population may account for some of the discrepancy. Study methods and genetic differences in the populations may also contribute to the different results.

Our data underscore the notion that, in children, the use of BMI as an indicator of adiposity has an important limitation because of individual variation in growth rates and maturity levels.29 Recent studies report that BMI may not accurately reflect adiposity in children, particularly among male adolescents and children of lower BMI.6,30 In this study, we found that, among boys ≥12 years of age, %BF and %TF remained relatively constant over a range of BMI until a higher BMI (>20 kg/m² for %BF and >16 kg/m² for %TF) was reached. A recent study in white males found similar results. It suggested that changes in BMI in males of lower BMI percentiles may occur without appreciable changes in adiposity.30 On the other hand, these observations support the use of high BMI percentile cutoff points (eg, 85th or 95th percentile) to identify children at risk for obesity. Furthermore, the distribution of fat seems to play a central role in the relation between obesity and blood pressure. Individuals with more visceral or android distribution of body fat are at higher risks for diabetes and heart disease.31 Therefore, additional studies are needed to understand the role of BMI as compared with direct measures of body fat in predicting cardiovascular risk in lean and overweight populations.

WC is a measure of central obesity that is easily obtained in medical practice. In addition, it is a characteristic of the metabolic syndrome.32 Lee et al33 reported that WC is significantly associated with total fat and insulin sensitivity and is an independent predictor of insulin resistance in black and white youths aged 8 to 17 years. Al-Sendi et al34 reported that WC is useful in identifying children (12–17 years of age) at risk of developing hypertension. In this study, there were gender differences in the correlations of WC with %BF and %TF in puberty such that WC did not accurately reflect body fat and fat distribution in pubertal boys. This may be because in a lean population of growing children, WC does not reflect changes in central deposition of body fat.

Finally, this is a relatively lean rural Chinese population. Using the growth chart from the Centers for Disease Control and Prevention,35 we plotted the Chinese children’s median BMI for specific age and gender (data not shown). We found that, on average, the Chinese children had lower BMI. The average difference in BMI between US and Chinese children is 1.5 kg/m² (boys) and 1.8 kg/m² (girls) at 6 to 11 years of age and 2.2 kg/m² (boys) and 2.0 kg/m² (girls) at 12 to 18 years of age. Although we are not certain whether the same relationships between BMI and adiposity measures that we found in the Chinese children will hold in US children, especially among overweight and obese children, our study underscored the notion that BMI and WC may not be accurate surrogates for total and trunk fat, especially in adolescent boys. It is important to understand such a relationship in each population so that clinical and research assessment of adiposity and associated health risk could be accurately and efficiently conducted.

CONCLUSIONS

In this relatively lean rural Chinese population, BMI and WC do not accurately reflect body composition as measured by DEXA in boys ≥12 years of age. In pubertal boys, body fat (except %TF) from DEXA did not show a parallel increase with BMI until the higher end of BMI. In contrast, both BMI and WC correlated well with %BF measured by DEXA in boys 6 to 11 years and in girls 6 to 18 years of age. The results demonstrate that it is essential to consider pubertal stage and gender when using BMI and WC to describe adiposity. These findings are important to consider in both research and clinical settings to accurately identify individuals with increased adiposity who are at risk for cardiovascular diseases or other adiposity-related morbidities.

ACKNOWLEDGMENTS

This study was supported in part by National Institutes of Health grant R01 HD049059 and by the Food Allergy Project.

We thank the team investigators who have provided
valuable input into the study, including Drs Katherine Kaufer Christoffel, Wendy Brickman, and Donald Zimmerman.

REFERENCES


Growth and Growth Hormone Therapy in Subjects With Mulibrey Nanism

Niklas Karlberg, MDa,b, Hannu Jalanko, MD, PhDa,b, Marita Lipsanen-Nyman, MD, PhDa

aDepartment of Pediatric Endocrinology, Hospital for Children and Adolescents, University of Helsinki, Helsinki, Finland; bBiomedicum Helsinki, Helsinki, Finland

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

ABSTRACT

OBJECTIVES. Mulibrey nanism is a monogenic disorder with prenatal-onset growth restriction, mild dysmorphic features, and a strong tendency for insulin resistance but no major neurologic handicap. Growth hormone therapy has been shown to promote short-term growth in children born small for gestational age, but the experience with long-term therapy is insufficient. Growth in patients with mulibrey nanism has not been analyzed previously in detail.

METHODS. We evaluated the natural growth pattern and long-term impact of growth hormone treatment in the largest cohort of subjects with mulibrey nanism to date. The study included 72 living subjects followed up to 30 years. Thirty (18 female) were treated with recombinant human growth hormone for a median period of 5.7 years. Patients were reviewed at baseline and every 6 to 12 months during the therapy. Evaluation included assessment of height, weight, and pubertal status and laboratory analyses. Glucose metabolism was evaluated by oral glucose-tolerance test.

RESULTS. The patients were born small for gestational age with immature craniofacial features. They experienced a continuous deceleration in height (median decrement of 1.1 SDS) and weight for height (median reduction of 17%) in infancy followed by an incomplete catch-up growth lasting up to school age. The final adult height averaged 136 cm in girls and 150 cm in boys. Growth hormone treatment improved the prepubertal growth but had only little impact on adult height (+5 cm). The treated subjects showed earlier bone maturation and growth arrest but not a significant increase in insulin resistance. On the contrary, the subjects who were treated with growth hormone were slimmer and had less metabolic syndrome as young adults.

CONCLUSIONS. The patients with mulibrey nanism showed a distinct postnatal growth pattern. The growth hormone treatment was safe and induced a good short-term effect, but the impact on the adult height remained modest.
MULTIBREY (MUSCLE-LIVER-EYE) NANISM (MUL; Online Mendelian Inheritance in Man No. 253250) is a rare autosomal recessive disorder with growth restriction beginning in utero.1,2 Today some 130 patients are known worldwide, 88 of them from Finland. Sporadic cases have been reported from different ethnic groups all over the world.2 MUL is caused by mutations in the TRIM37 gene located on chromosome 17q22-q23.3 It encodes for TRIM37 protein, which is a member of the tripartite motif protein family (TRIM; composed of RING, B-box, and coiled-coil domains).4 TRIM37 is expressed in several tissues,5,6 has been localized to peroxisomes,4 and possesses E3 ubiquitin-ligase activity as several other RING proteins.7,8 Twelve different mutations have been reported in MUL patients, with 3 of them present in the Finnish patients with MUL: Fin-major (c.493-2A>G), Fin-minor (c.2212delG), and a new c.227T mutation.3,7,9 The dysmorphic features in MUL patients form a distinct entity with characteristic craniofacial features and constitutional gracility.2 Other typical findings are fibrous dysplasia of long bones, hepatomegality, restrictive heart disease, cutaneous naevi flammei, yellowish dots in ocular fundi, and a J-shaped sella turcica.1,2 The patients are neurologically normal, but often mild muscular hypotonicity and slight delay in motor and speech development is evident.2 Feeding difficulties and respiratory tract infections, including severe pneumonias, are quite common in infancy. Congestive heart failure appears in ~10% of infants.2,10 An interesting recent finding is that patients with MUL show a dramatic change in glucose and lipid metabolism with age. Although the infants and children have a tendency for low glucose and insulin levels, a great majority developed severe insulin resistance and a full-blown metabolic syndrome between the ages of 11 and 20 years.11

Subjects with MUL form a homogenous group of small for gestational age (SGA) children. Growth hormone (GH) therapy has been shown to promote short-term growth both in SGA children and in children with a number of dysmorphic disorders, including Silver-Russell syndrome, but the experience with long-term therapy is still insufficient.1,2,4 Here we report the natural growth pattern of patients with MUL and their responsiveness to GH treatment. The results show a good short-term effect of GH in prepubertal children but only a modest impact on the adult height. Although subjects with MUL have a strong tendency for insulin resistance,11 GH seemed to have positive long-term effects on body weight and glucose metabolism.

PATIENTS AND METHODS

Patients

Growth data on 72 patients (40 female) from 68 families were analyzed. All fulfilled the clinical diagnostic criteria of MUL.1,2 The diagnosis was also confirmed by genetic testing: 69 were homozygous for the Fin-major mutation, 2 were compound heterozygotes for the Fin-major and Fin-minor mutations, and 1 subject had the c.227T>C/Fin-major genotype. In infancy (0–24 months), upper respiratory tract infections and pneumonias were common problems, and half of the patients (38 of 72) exhibited some feeding difficulties, warranting nasogastic feeding or percutaneous gastrostomy (n = 6) in 21 of them. The feeding difficulty was classified as severe in 5 children because of prolonged and extremely poor and slow feeding, including fatigue and vomiting or total food neglect with no attempt for oral feeding. Congestive heart failure was diagnosed in 9 infants (13%) at a median age of 1.1 years.

Patient Follow-up

Since the early 1970s, clinical care of the Finnish patients with MUL has been centered in our institution. The patients have had physical examinations at 6- to 12-month intervals during childhood and puberty. Standing heights were measured with a stadiometer to the nearest millimeter, and the mean of 3 measurements was used. The weight was measured with an ordinary scale to the nearest 100 g. Height and weight measurements were made by the clinician or a trained nurse. When needed, growth data from birth records, child welfare centers, schools, and local hospitals were collected. Birth data were available from all 72 of the patients. A radiograph of the palm and wrist was obtained for evaluation of the bone age (BA). At baseline and annually thereafter BA was determined by a single observer (Dr Lipsanen-Nyman) according to the method described by Greulich and Pyle.15

GH Therapy

Thirty subjects (18 female) have been treated with recombinant human GH since the early 1990s. Fourteen of them are still on GH, and the duration of treatment ranges from 1.8 to 13.4 years (median: 5.7 years). The GH dose has been 0.035 mg/kg per day, and it has been adjusted to the patients’ weight during the treatment. Decision on GH treatment was made on clinical grounds and was mainly based on auxological data. The patients were reviewed at baseline and at 3, 6, 9, and 12 months after initiation of the treatment and subsequently every 6 months.

Laboratory Examinations

Blood glucose, serum insulin, lipids, leptin, and insulin-like growth factor (IGF) 1 were measured annually since 1999 after an overnight fast. Serum IGF-1 levels were measured by radioimmunoassay (Incstar, Stillwater, MN), and serum leptin levels were assessed by specific radioimmunoassay (Linco Research Inc, St Louis, MO). A standard arginine or insulin-arginine tolerance test was used to evaluate the GH secretion. Serum GH con-
centrations were measured by monoclonal immunoradiometric assay (CIS Bio International, Gif-sur-Yvette, France) before the year 1995 and thereafter by Auto-Delfia time-resolved immunofluorometric assay (Perkin-Elmer, Wallac, Turku, Finland). A peak serum GH response of >10 μg/L was considered normal, and a value 5 to 10 μg/L indicated partial GH deficiency. The patient was considered deficient in GH if the peak serum level was <5 μg/L in 2 separate tests. In patients receiving GH, blood counts and serum concentrations for liver, thyroid, and kidney function and glycohemoglobin A1c were followed as safety parameters. A 3-hour oral glucose-tolerance test (glucose load: 1.75 g/kg; maximum: 75 g) was performed on a subset of the patients, and glucose tolerance and insulin sensitivity were interpreted as described before.11

Data Analyses
Finnish growth standards were used for height and weight analysis.17 Height SD score (hSDS) was calculated for calendar age. Final hSDS was calculated from the standards of the Finns at the age of 18 years.

The weight of the patients was defined as the percentage of deviation from the age-specific median weight for height (WFH).18 Adult weights were also expressed as the BMI. Birth length, weight, and head circumference were expressed as SDSs according to the Finnish standards. The birth length and weight of patients born preterm (before 38 weeks of gestation) were extrapolated to term by using Finnish standards of prenatal growth.19 The onset of puberty was defined as stage P2G2 in boys and a stable breast stage of M2 in girls.20

Growth data were correlated to metabolic and endocrine parameters as described before.11 Wilcoxon test was performed for comparisons between the GH-treated and untreated patient groups. P values <.05 were defined as statistically significant. Metabolic and endocrine parameters were correlated with the hSDS increment during the preceding year, and the explanation rate was expressed as an R² value.

The ethics committee of the Hospital for Children and Adolescents, University of Helsinki, approved the study. All of the patients or their guardians gave informed consent.

RESULTS

Size at Birth
At birth, the infants with MUL were both short and light. The median birth length and weight adjusted to 40 weeks of gestation were 44.8 cm (median hSDS: −3.0) and 2300 g (median SDS: −3.0) for the girls and 45.0 cm (median hSDS: −2.8) and 2350 g (median SDS: −2.9) for the boys, respectively. The median occipitofrontal head circumference SDS was −0.5 (range: −0.9 to 0.8), indicating spared cranial growth with macrocephaly relative to birth length.

Height in Untreated Patients With MUL
The growth failure progressed in infancy with a median hSDS decrement of 1.1 from birth to 2.0 years of age (Fig 1). In 9 children with congestive heart failure and 5 with severe prolonged feeding difficulties, the linear growth decelerated even further. Their median hSDS at 2.0 years of age was −5.1 (range: −5.9 to −3.8) and −5.4 (range: −4.3 to −7.4), respectively, as compared with hSDS of −4.4 in children only mildly affected with these problems. Two children were born severely premature at gestation week 32, with a birth length SDS of −6.4 and −4.0, respectively. Their postnatal growth was most severely affected so that their hSDS at 2.0 years of age was −7.8 and −7.3.

Infant growth deceleration was followed by a sustained period of spontaneous incomplete catch-up growth, which lasted up to school age (Fig 1). From 7 to 8 years onward, the median hSDS remained fairly stable in most patients. At the onset of puberty, the median hSDS was −3.6 (range: −6.3 to −1.0; Table 1 and Fig 1). The pubertal growth spurt was weak or absent (Fig 1). Twenty-five of the patients (13 female) have reached their adult height. The final adult height was, on average, 136 cm (range: 130–155 cm; median hSDS: −5.1) in female subjects and 150 cm (range: 147–162 cm; median hSDS: −4.1) in male subjects.

Height in Patients With MUL Receiving GH Therapy
Thirty patients (18 female) have received GH therapy on average for 5.7 years (range: 1.8–13.4 years). At commencement of GH therapy, the median age was 4.4 years (range: 1.6–8.9 years), and the median hSDS was −4.7 (range: −7.8 to −2.0). The GH treatment improved growth temporarily so that the growth velocity reached its peak 12 to 18 months after the start (Figs 1 and 2 and Table 1). During the first year, the increment was 1.0 SDS on average, and by the onset of puberty, it was 1.8 SDS (range: 0.2–2.7 SDS). The hSDS was at pubertal onset 0.7 greater in the GH-treated compared with the untreated children (P < .02). During puberty, the median hSDS remained stable at −2.9 (Table 1). Two children with prolonged severe feeding difficulties and 3 with congestive heart failure were treated with GH. Three of them responded very poorly to the therapy, and the treatment was discontinued after 1.8, 2.0, and 2.4 years. The GH treatment still continues in 2 subjects, but the response has been modest. Both of the 2 children born severely premature received GH, but their responses have so far been poor.

By now, 16 patients (8 female) receiving GH have reached their adult height with a median of 142 cm (range: 137–154 cm) and 155 cm (range: 150–163 cm) in female and male subjects, respectively (Fig 2).
corresponding median adult hSDS according to the Finnish standards was 4.2 in the female subjects and 3.6 in the male subjects (Fig 1 and Table 1). Thus, the difference in the final hSDS between the treated and untreated patients was 0.6 (5 cm; P < 0.03; Fig 1). The female subjects seemed to benefit slightly more (by 0.4 hSDS) than the male subjects (P < 0.05). The pubertal growth was poor (Fig 1). The GH-treated female and male subjects reached their final height 1.1 and 0.6 years earlier than their untreated counterparts (P < 0.005 and P < 0.03, respectively). No correlation was found between the adult height and the age at start of the GH treatment, duration of the treatment, height at onset of puberty, GH peak during provocation, or the serum IGF-1 levels before commencement of the GH treatment.

**WFH in GH-Treated and Untreated Subjects**

Weight gain during infancy was very poor in all of the children, and during the 2 first years there was a con-

---

**TABLE 1** Pubertal Growth in 41 Postpubertal Patients With MUL

<table>
<thead>
<tr>
<th>Variable</th>
<th>Natural Growth</th>
<th>GH Therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pubertal stage P2G2/M2, y</td>
<td>12.5 9.2 to 15</td>
<td>11.6 8.8 to 15</td>
</tr>
<tr>
<td>Menarche, y</td>
<td>15 10.5 to 17</td>
<td>12 10 to 15.5</td>
</tr>
<tr>
<td>Height at P2G2/M2, SDS</td>
<td>-3.6 -6.3 to 1.4</td>
<td>-2.9 -4.2 to 0.9</td>
</tr>
<tr>
<td>WFH at P2G2/M2, %</td>
<td>-14 -22 to 20</td>
<td>-14.5 -24 to 11</td>
</tr>
<tr>
<td>Height 2 y after P2G2/M2, SDS</td>
<td>-3.5 -5.8 to 1.3</td>
<td>-2.9 -5.3 to 1.8</td>
</tr>
<tr>
<td>WFH 2 y after P2G2/M2, %</td>
<td>-3 -21 to 58</td>
<td>-8 -21 to 20</td>
</tr>
<tr>
<td>Adult height, SDS</td>
<td>-4.6 -6.7 to 2.2</td>
<td>-4.0 -4.9 to 2.1</td>
</tr>
<tr>
<td>WFH at adult height, %</td>
<td>20 -2 to 75</td>
<td>3 -20 to 34</td>
</tr>
</tbody>
</table>

Sixteen patients (8 female) were treated with GH, and 25 had no growth-promoting treatment. GH indicates recombinant human GH at a dose of 0.035 mg/kg per day.

*Values are from 21 female subjects: 13 untreated and 8 treated with GH.*
continuous deceleration of the WFH (Fig 1). Children remained thin until puberty when WFH started to increase, more rapidly in the girls (Fig 1).

After commencement of GH therapy the median WFH increased. At onset of puberty, no difference in WFH was observed between the GH-treated and untreated subjects (Table 1). However, when reaching the adult height, the subjects treated with GH had remained slimmer than the untreated patients. Their median WFH was 0% (−20% to 24%) in the male subjects and 16% (−3% to 34%) in the female subjects, whereas the corresponding values in the untreated patients were 15% (−2% to 25%) and 25% (3% to 75%), respectively (P < .005; Table 1 and Fig 1).

**BA and Puberty in GH-Treated and Untreated Subjects**

The BA in young children was clearly delayed relative to chronologic age with a BA/chronologic age of 0.85 at 2 years. No significant difference in BA was found between the GH-treated and untreated subjects up to the age of 8 years (Fig 1). Thereafter, the bone maturation accelerated, particularly in the GH-treated patients, so that a significant difference (1.09 vs 0.92) in BA was evident at the age of 10 years (P < .01). At the onset of puberty, the ratio of BA and chronological age was 1.06 and 0.92 in the GH-treated and untreated patients, respectively (P < .03). The difference remained significant throughout puberty so that this ratio at the age of 14 years was 1.05 and 0.95 in the treated and untreated patients, respectively (P < .03; Figs 1 and 3).

There was a wide individual variation in the onset of puberty with a median age of 12.5 and 13.0 years for boys and girls, respectively (Table 1). No significant difference was noted between the GH-treated and untreated patients. However, the duration from stage 2 breast development to menarche was shorter, and menarche occurred ~3.0 years earlier in the GH-treated compared with untreated girls (Table 1).

**GH Production in Children With MUL**

The GH production was assessed by the arginine or arginine-insulin stimulation tests in 38 children with MUL at a median age of 5.1 years (range: 2.0 –15 years). The response was normal (~10 g/L) in 26% (10 of 38), subnormal (range: 5–10 g/L) in 58% (22 of 38), and low (~5 g/L) in 16% (6 of 38). The distribution was similar in those 26 individuals who were treated with GH after the tests (27% normal, 62% subnormal, and 11% low). Serum IGF-1 levels were evaluated in 26 prepubertal children who had not received GH treatment. All had normal serum concentrations of IGF-1, but half of them had an IGF-1 value in the lowest quartile of the reference range.

**Growth and Glucose Metabolism in Children With MUL**

Data on glucose metabolism, including an oral glucose-tolerance test, were available from 13 children (median age: 6.4 years; range: 3.4–9.7 years) who had received GH during the preceding year (6 currently on GH) and 17 without any growth promoting therapy (median age: 7.1 years; range: 3.2–10.6 years). In both groups, the patients with the greatest hSDS increment during the preceding year presented the highest levels of fasting serum insulin (R² = 0.69 and 0.62), postload peak in-
sulin (R² = 0.72 and 0.72), serum IGF-1 (R² = 0.54 and 0.58), and serum leptin (R² = 0.57 and 0.58; Fig 4). In all 30 of the children, serum fasting insulin and postload peak insulin levels correlated well with the serum IGF-1 concentration with explanation rates (R² value) of 0.66 and 0.71, respectively (Fig 4). Overall, the GH-treated children had slightly higher serum fasting insulin (median: 9 vs 5 mU/L; P < .05) and postload peak insulin (median: 152 vs 89 mU/L; P = .08, not significant) concentrations compared with the untreated children. The same was noted in pubertal female subjects. The difference in fasting (15 vs 11 mU/L) or postload peak insulin concentrations (200 vs 173 mU/L), however, did not reach statistical significance.

FIGURE 3
A, Height and BA in a girl treated with GH for 10.6 years. BA is delayed but advances by the age of 8 years. B, The same female subject at 6 months and 2.0, 6.0, and 20 years. The change from a thin child to an overweight adult (WFH 18%) is obvious.

FIGURE 4
Metabolic and endocrine parameters correlated with hSDS increment during the preceding year in 13 prepubertal children on GH and 17 children without growth-promoting therapy. In both groups, the greatest hSDS increment predicted the highest serum fasting (A and B) and postload oral glucose-tolerance test for peak insulin (C and D), IGF-1 (E and F), and leptin (G and H).
Glucose Metabolism in GH-Treated and Untreated Young Adults With MUL

Glucose metabolism was analyzed in 12 adults (median age: 24.2 years) who had been treated with GH in childhood and in 11 untreated subjects (median age: 23.9 years). Patients in the treatment group were lighter (WFH 0% vs 15%; \( P < .04 \)), had a lower BMI (17 vs 21 kg/m\(^2\); \( P < .04 \)), and showed lower fasting blood glucose (4.4 vs 5.0 mmol/L; \( P < .05 \)) than the untreated subjects. Moreover, their blood pressure (median: 120/76 vs 133/83 mm Hg), total serum cholesterol (median: 4.0 vs 5.2 mmol/L), and postload peak serum insulin (192 vs 250 mU/L) were lower than in untreated subjects, although the comparison did not reach statistical significance (Table 3). Also, the frequency of metabolic syndrome was twice as high (64% vs 33%) in the GH-treated compared with the untreated subjects.

**DISCUSSION**

Of all newborns, 2.5% are born SGA, and 10% to 15% of them lack postnatal catch-up growth and remain persistently short.\(^2\) Many of them have chromosome abnormalities, monogenic disorders, or familial or sporadic syndromes.\(^2\) We report here the natural growth pattern and impact of long-term GH treatment in a monogenic disorder, MUL. To our knowledge, this study provides the longest follow-up time of the GH therapy in subjects who represent a homogenous subgroup of children with prenatal growth restriction.

Our results clearly show that patients with MUL are born SGA and not only fail in early postnatal catch-up growth but also experience a continuous deceleration both in height and weight development through infancy (≤24 months). This wasting is followed by a subsequent spontaneous but incomplete catch-up growth until the age of 7 to 8 years. This growth pattern and the early catch-down in height and weight resembles those of infants born very preterm (at weeks 23–25 of gestation) experiencing postnatal extrauterine growth restriction.\(^2\)\(^3\)\(^\#\) Growth in preterm infants decelerates as they become exposed to extrauterine life, and their energy expenditure shifts to promote survival rather than growth.\(^2\)\(^3\) Infants with MUL born at term not only grow as premature infants, they also present many premature features, such as relative macrocephaly, wide skull sutures and prominent forehead (Fig 3), and an immature craniofacial skeleton: the peculiar shape of sella turcica in MUL resembles the shape of the bony sphenoid before the cartilaginous dorsum sellae has ossified.\(^2\)\(^6\) Also, the shape of the skull and the size relationship between the skull and the face are in line with an idea of delayed prenatal and postnatal maturation in MUL.\(^2\)\(^,\)\(^2\)\(^6\) The fact that half of the infants with MUL have serious feeding difficulties and are prone to severe pneumonias proposes immaturity as well, because similar problems are common among preterm infants with immature lungs and bowel function.\(^2\)\(^5\)\(^,\)\(^2\)\(^7\) Although the frequency of feeding difficulties in patients with MUL is strikingly high, it

### TABLE 3  Glucose and Fat Metabolism in Young Adults With MUL

<table>
<thead>
<tr>
<th>Variable</th>
<th>Natural Growth</th>
<th>GH Therapy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>Range</td>
</tr>
<tr>
<td><strong>Anthropometrics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adult height, SDS</td>
<td>-4.4</td>
<td>-5.9 to 1.4</td>
</tr>
<tr>
<td>WFH, %</td>
<td>15</td>
<td>-2 to 67</td>
</tr>
<tr>
<td>BMI, kg/m(^2)</td>
<td>20.5</td>
<td>17.2 to 30.9</td>
</tr>
<tr>
<td>SBP, mm Hg</td>
<td>133</td>
<td>110 to 168</td>
</tr>
<tr>
<td>DBP, mm Hg</td>
<td>83</td>
<td>72 to 118</td>
</tr>
<tr>
<td><strong>Glucose metabolism</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS-glucose, mmol/L</td>
<td>5.0</td>
<td>4.1 to 6.8</td>
</tr>
<tr>
<td>OGTT 2-h glucose, mmol/L</td>
<td>9.1</td>
<td>6.8 to 13.2</td>
</tr>
<tr>
<td>B-GHba1cM, %</td>
<td>5.3</td>
<td>4.6 to 6.7</td>
</tr>
<tr>
<td>FS-insulin, mU/L</td>
<td>23</td>
<td>6 to 56</td>
</tr>
<tr>
<td>Postload peak insulin, mU/L</td>
<td>250</td>
<td>150 to 1370</td>
</tr>
<tr>
<td><strong>Fat metabolism</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FS-cholesterol, mmol/L</td>
<td>5.2</td>
<td>3.3 to 6.6</td>
</tr>
<tr>
<td>FS-LDL, mmol/L</td>
<td>2.6</td>
<td>0.9 to 4.3</td>
</tr>
<tr>
<td>FS-HDL, mmol/L</td>
<td>1.0</td>
<td>0.7 to 1.3</td>
</tr>
<tr>
<td>FS-triglycerides, mmol/L</td>
<td>1.9</td>
<td>0.7 to 4.7</td>
</tr>
<tr>
<td>S-leptin, µg/L</td>
<td>17.7</td>
<td>3.8 to 39.9</td>
</tr>
<tr>
<td>S-uric acid, µmol/L</td>
<td>429</td>
<td>335 to 495</td>
</tr>
<tr>
<td><strong>Insulin sensitivity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FGIR</td>
<td>0.23</td>
<td>1.9 to 0.09</td>
</tr>
<tr>
<td>Stumvoll</td>
<td>0.090</td>
<td>0.13 to 0.03</td>
</tr>
</tbody>
</table>

Data on 12 subjects (5 female) treated earlier with GH (median age: 23.9 years) were compared with 11 untreated subjects (5 female; median age: 24.2 years). SBP indicates systolic blood pressure; DBP, diastolic blood pressure; LDL, low-density lipoprotein; HDL, high-density lipoprotein; FS, fasting serum; OGTT, oral glucose-tolerance test; FGIR, fasting glucose insulin ratio; Stumvoll, whole-body insulin sensitivity index calculated according to the formula developed by Stumvoll.\(^1\)
seems that only severe and prolonged feeding difficulties had a negative impact on the linear growth.

Most patients with MUL had at least partial GH deficiency and showed subnormal or low-normal levels of serum IGF-1. Evidence of disturbances in the GH-IGF axis, such as GH deficiency and resistance, has been reported previously in short children born SGA.\textsuperscript{28,29} IGF-1 and IGF-2 are major regulators of prenatal and postnatal growth, and reduced IGF-1 levels have been observed in children who are born SGA.\textsuperscript{29,30} Also, gene-knockout studies have shown a major impact of IGF-1 on prenatal and postnatal growth.\textsuperscript{31} Recent observations of variation in methylation patterns regulating imprinting and expression of IGF-2 in Silver-Russell syndrome further highlight the role of growth factors in IUGR.\textsuperscript{31} The precise role of the GH-IGF axis in the poor growth of subjects with MUL remains to be elucidated.

The long-term GH treatment had little impact on the adult height of subjects with MUL (hSDS increment of 0.6 = 5 cm), which is in line with results reported previously in patients with Silver-Russell syndrome and skeletal dysplasias.\textsuperscript{31,32} The gain in height was mainly achieved before the onset of puberty, as has been seen in previous studies on SGA.\textsuperscript{31} A substantial catch-up growth was observed, especially in the first 2 years, and, at the time of discontinuation of the GH therapy, there was a median hSDS increment of 1.9 compared with baseline data. Bone maturation and growth arrest, however, occurred early in patients receiving GH, which explains why the adult height in the treated patients still remained very poor as compared with the general population. The subjects treated with GH were slightly smaller at 2 years of age (median hSDS: −4.4 vs −4.7), which may cause a small bias. The adult height correlated poorly with the parental target height, and no predictive factor for the growth response was found. Five patients (3 on GH) with only minor heart and neonatal problems reached closest to their target height, suggesting that the overall severity of the disease influences the linear growth. However, even these subjects remained between −1.5 and −3.2 hSDS below their target height.

It is interesting to note that most MUL patients develop insulin resistance in childhood and metabolic syndrome with severe insulin resistance after puberty.\textsuperscript{11} Because GH is an insulin antagonist and may induce hyperinsulinemia, we were especially concerned about the glucose metabolism in the GH-treated patients with MUL. Importantly, the preexisting insulin resistance was not significantly increased by the GH therapy in prepubertal or pubertal children. Moreover, the adult patients who had received GH had lowered WFH and BMI, and their deterioration of the glucose metabolism seemed less severe as compared with untreated subjects with MUL. Whether this holds when a larger number of patients with MUL has reached the adulthood remains to be seen. However, favorable changes in body composition, BMI, and the lipid atherogenic index have been presented in large cohorts of SGA children on GH therapy.\textsuperscript{34}

It is also interesting to note that the greatest increment in height was observed in those children with the highest fasting and postload peak serum insulin concentrations, suggesting that insulin might improve the spontaneous childhood growth and accelerate the short-term catch-up growth after commencement of GH therapy. Indeed, insulin is a well-known growth-promoting factor both prenatally and postnataally. It regulates the IGF-1 concentration by facilitating the binding of GH to its receptor, stimulates the production of IGF-1, and increases bioavailability of IGF-1 by suppressing the hepatic synthesis of IGF binding protein.\textsuperscript{35,36} Moreover, in a recent study involving non-GH-treated SGA girls, a rapid progression in pubertal growth tempo and progression to menarche driven by insulin was observed.\textsuperscript{37,38} Our results are in line with this, suggesting that the aggrivated levels of insulin might underlie the faster pubertal tempo to menarche seen in female subjects with MUL on GH treatment. However, although female subjects with MUL have spontaneous puberty with menarche, they present early irregularity of menstrual periods with subsequent ovarian failure.\textsuperscript{39} Because pubertal growth is driven primarily by estrogen,\textsuperscript{40} it is possible that sex steroids have an impact on the poor pubertal growth in patients with MUL.

The TRIM37 mutations in the Finnish patients with MUL result in a nonfunctional TRIM37 protein. TRIM37 has been localized to peroxisomes in cell cultures and found to act as an E3 ubiquitin ligase.\textsuperscript{41} The ubiquitin-proteasome pathway has been implicated in the pathogenesis of type 2 diabetes and growth failure.\textsuperscript{11,41} GH receptor is a key regulator of cellular metabolism and requires an active ubiquitination system for both endocytosis and degradation of GH receptor.\textsuperscript{42} Recently discovered mutations in the gene cullin 7 (CUL7) on human chromosome 6p21 have been identified to result in the autosomal recessive 3-M syndrome, with features resembling those of MUL, including severe prenatal and postnatal growth retardation, facial dysmorphism, large head circumference, and normal intelligence. It is interesting to note that the CUL7 is crucial in assembling an E3 ubiquitin-ligase complex and thereby promotes ubiquitination.\textsuperscript{43} Mutations in the CUL7 gene result in defective ubiquitination linking impaired ubiquitination to the pathogenesis of prenatal-onset growth failure in humans. MUL seems to be a novel example of such conditions.

**CONCLUSIONS**

Our study provides results on the GH treatment in a homogenous subgroup of SGA children failing to catch up. The children with MUL share several features with
the very preterm infants, and the data may bring information to the debate on the GH treatment in very preterm children experiencing extraterine growth restriction.\textsuperscript{23,25} The GH therapy in the subjects with MUL was safe and improved the growth in the short-term. However, the impact of GH therapy on the adult height was modest compared with the short-term response, which should be kept in mind when this mode of therapy is planned. Whether the same situation is noted in other growth-deficiency syndromes or in very preterm children remains to be seen.

ACKNOWLEDGMENTS

This study was supported by the Finnish Foundation for Pediatric Research, Finska Läkaresällskapet, Finnish State Grant TYH 3304, the Finnish Academy, the Sigrid Juselius Foundation, the Wilhelm and Else Stockmann Foundation, and the K. Albin Johansson Foundation.

We thank the families of our patients and express our gratitude to the patients’ local pediatricians for excellent cowork. We express special gratitude to Prof Jaakko Perheentupa for his pivotal early role in the diagnosis and treatment of patients with Mulibrey nanism.

REFERENCES

14. de Zegher F, Hokken-Koelega A. Growth hormone therapy for children born small for gestational age: height gain is less dose dependent over the long term than over the short term. \textit{Pediatrics.} 2005;115(4). Available at: www.pediatrics.org/cgi/content/full/115/4/e458
30. Murphy VE, Smith R, Giles WB, Clifton VL. Endocrine regu-


38. Ibanez L, Ong K, Valls C, Marcos MV, Dunger DB, De Zegher F. Metformin treatment to prevent early puberty in girls with precocious pubarche. *J Clin Endocrinol Metab.* 2006;91:2068–2073


Parental Ability to Discriminate the Weight Status of Children: Results of a Survey

Jeannie S. Huang, MD, MPH\textsuperscript{a,b}, Karen Becerra, DDS\textsuperscript{c}, Thaghar Oda, MD\textsuperscript{d}, Edward Walker, BS\textsuperscript{c}, Ronghui Xu, PhD\textsuperscript{c}, Michael Donohue, PhD\textsuperscript{c}, Iris Chen\textsuperscript{a}, Violeta Curbelo, MD\textsuperscript{d}, Adam Breslow, MD\textsuperscript{d}

\textsuperscript{a}Division of Biostatistics and Bioinformatics and \textsuperscript{b}Department of Pediatrics, University of California, San Diego, California; \textsuperscript{c}Department of Gastroenterology, Children’s Hospital San Diego, San Diego, California; \textsuperscript{d}Children’s Primary Care Medical Group, San Diego, California

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

ABSTRACT

OBJECTIVES. In this study we aimed to explore parents’ weight perceptions of their children and of unrelated children.

METHODS. Parents of children \(\leq 18\) years of age who were attending pediatric clinics throughout San Diego County, California, were surveyed concerning their children’s weight status and the weight status of unrelated children in various age groups. Height and weight were measured, and weight status was determined for both the parent and child. The influence of various demographic variables on parents’ weight perceptions and the relationship between parents’ perceptions of weight of their children and parents’ perceptions of weight of unrelated children were evaluated. Multivariate regression modeling was applied to identify predictors of parents’ perceptions of weight of their own children.

RESULTS. Of 1098 parents surveyed, 87% were women, 74% were white, and 46% reported Hispanic ethnicity. Seventy percent of the parents surveyed were overweight or obese, and 39% of their children were at risk for overweight or overweight. Sixty-one percent of parents correctly identified their children’s weight status, and parents were able to correctly identify the weight status of unrelated children in 58% of reviewed photographs. Parents’ weight perceptions of their children were not related to their ability to determine the weight status of unrelated children or to their ideal weight selections among unrelated children. In a multivariate logistic regression analysis, parental ability to correctly assess their child’s weight status was associated with their child’s age and weight status.

CONCLUSIONS. Parents’ perceptions of their own children’s weight status are influenced by their children’s characteristics and do not seem to correspond with their weight perceptions of unrelated children. Parental recognition of weight issues in their offspring may be impeded by their inability to apply criteria used to ascertain the weight status of unrelated children to their own children.
OveWEIGHT IN CHILDREN is a national health concern with increasing prevalence rates in children of all ages over the past 2 decades.¹ Health morbidities associated with overweight in children and adolescents include type 2 diabetes,² obstructive sleep apnea,³ liver disease,⁴ hypertension and increased cardiovascular risk,⁵ and poor quality of life.⁶

One frequently cited reason as to why childhood overweight is on the rise is the failure of parents to recognize the overweight status of their children.⁷–¹² Recognition rates of the heavier weight status of at risk for overweight (AROW) and/or overweight children by parents vary widely (from 6.2% to 66.7%⁷–¹⁵) and may reflect different populations and assessment methods. Certain child and parent characteristics, such as child weight status,⁷–¹³,¹⁵ child age,⁹,¹³ child sex,¹⁰,¹²,¹⁵ and parental education level,⁸ have been suggested to influence parental recognition of overweight in their children. Investigators have proposed that lack of parental recognition of the overweight status of their children may reflect a general inability of parents to distinguish abnormal from normal weight status because of the increasing prevalence of heavier body types in the general population.¹⁰,¹¹ However, data are lacking regarding whether parental perceptions of weight of unrelated children are similar to their perceptions of the weight of their own children.

This study evaluated parents’ perceptions of weight of their children and of unrelated children and explored the influence of various demographic variables on these weight perceptions. To assess parental ability to assess the weight status of unrelated children, we developed a photograph survey of various-aged children of different weights. Our a priori hypothesis was that parents’ perceptions of weight of their own children are positively associated with their perceptions of weight of unrelated children and reflect public perceptions of weight of children as others have suggested.¹⁰,¹¹

### PATIENTS AND METHODS

**Participants and Setting**

All of the aspects of the study were reviewed and approved by the University of California, San Diego, internal review board. Informed consent was obtained from parents before study performance, and assent was obtained in children ≥7 years of age. Parents were recruited at the time of their child’s appointment at local pediatricians’ offices in the greater San Diego area, representing various sociodemographic, racial, and ethnic populations. Parents with children aged 0 to 18 years old were eligible for participation. Parents of children photographed for the photograph questionnaire were not eligible for participation in the questionnaire study.

A total of 1461 parent-child pairs were surveyed for this questionnaire representing 1098 families and 1106 parents. Only data from 1 parent-child pair per family were entered into the current study analyses. In cases where ≥1 parent-child pair was examined in a given family (in 363 families), data from the child whose first name came last alphabetically were entered into the analysis. In cases where both father and mother participated (in 8 families), data from the father were entered. Ultimately, 1098 parent-child pairs were represented in the study population.

**Study Performance**

After study procedures were reviewed and informed consent was obtained, parents surveyed were presented 12 photographs of children in 4 age categories (infant, toddler [1–3 years], 4–6 years, and 10–12 years) in the order listed in Table 1 and asked to determine whether photographed children were underweight, normal weight, AROW, or overweight. Subsequent to this initial survey, parents were then presented all of the photographs within a particular age category and asked to select the photograph that best depicted their ideal weight type for that age category. Parents were also asked to give an opinion regarding the weight status of their own child according to the weight categorizations of underweight, normal weight, AROW, and overweight. Demographic information was collected. Surveys were performed in both English and Spanish according to parental preference.

Heights and weights were measured in both parent and child, and BMI was calculated. Parent weight cate-
gory was assigned according to National Center for Health Statistics BMI categorical definitions: \( x < 18.5 \) kg/m\(^2\), underweight; \( 18.5 \leq x < 25 \) kg/m\(^2\), normal weight; \( 25 \leq x < 30 \) kg/m\(^2\), overweight; and \( x \geq 30 \) kg/m\(^2\), obese. Weight status for children \( \geq 2 \) years of age was defined according BMI for age and sex percentile definitions of underweight \( (x \leq 5\%)\), normal weight \( (5\% < x < 85\%)\), AROW \( (85\% \leq x < 95\%)\), and overweight \( (x \geq 95\%)\), as defined by the Maternal and Child Health Bureau, Health Resources and Services Administration, the Department of Health and Human Services.17 Children \( < 2 \) years of age were similarly assigned weight status on the basis of sex-specific weight-for-height percentiles and the following definitions: underweight \( (x \leq 5\%)\), normal weight \( (5\% < x < 85\%)\), AROW \( (85\% \leq x < 95\%)\), and overweight \( (x \geq 95\%)\).

Photograph Questionnaire
Children of various weights and ages were photographed in both the anteroposterior and lateral positions (Fig 1). Informed consent was obtained from parents and assent obtained from children \( \geq 7 \) years of age before photographing the pictured children in accordance with University of California internal review board requirements. Photographed children’s weight categories were determined according to BMI-for-age and sex percentiles \( (\geq 2 \text{ years})\) and weight-for-height age and sex percentiles \( (< 2 \text{ years})\) as described above. The age categories represented were as follows: infant \( (<1 \text{ year})\), toddler \( (1–3 \text{ years})\), child \( (4–6 \text{ years})\), and older child \( (10–13 \text{ years})\). Infants were photographed in the seated position in diapers. Older children were photographed standing in their undergarments. Anonymity of photographed patients was preserved via superimposing a black bar over identifying facial features. In each category, 3 photographs representing the 4 different weight categories were selected. For “normal weight,” we selected children in the midrange of the normal spectrum \( (\text{ie}, 25\%–75\%)\) to better represent the typical normal child. Table 1 lists the demographics and weight status of the child photographed according to the order of presentation in the survey. The presentation order of the photographs was determined randomly.

Definition of Variables
For surveyed parents, racial response categories included: white, black, Asian, or other; for statistical analyses, these groups were collapsed according to white or non-white origin. Other demographic variables were collapsed into dichotomous categories, including the following: ethnicity (Hispanic versus non-Hispanic), sex (male versus female), home language (English [includ-
ing monolingual and bilingual English-speaking homes] versus other), family size (1 vs >1 child), parent education status (any college or more versus high school education or less), income (poverty versus above poverty), number of family generations in the United States (≥2 vs <2), participant child age (≥4 vs <4 years), parent age (≥30 vs <30 years), and weight status (AROW and overweight versus normal weight and underweight). Poverty level was defined according to the US Census Bureau 2005 statistics for families.18

Characteristics of photographed children evaluated for these analyses were represented as follows: child sex (male versus female), child ethnicity (Hispanic versus non-Hispanic), child race (white versus nonwhite), child age (infant or toddler versus 4–6 or 10–14 years), and child weight status (AROW or overweight versus normal weight or underweight).

Parent responses for the photograph survey and for weight assignments for their own children were coded as correct or incorrect. Parental ability to correctly identify the weight status of unrelated children was represented as the number of photographed unrelated children for whom weight status was correctly identified (of 12). Parental ability to correctly identify the weight status of AROW or overweight unrelated children was represented as the number of photographed AROW or overweight children correctly identified (of 7). Parent-selected weight ideals in each age category were coded as AROW or overweight versus normal weight or underweight.

Statistical Analyses
Demographic data were summarized using descriptive statistics. Univariate analyses of parental ability to correctly identify unrelated children’s weight status by selected parent factors were performed using the Wilcoxon rank sum test. To assess the association between unrelated child characteristics and parental ability to correctly identify the weight status of unrelated children, the generalized estimating equation approach was used because of correlated outcome data (because each parent rated all of the 12 unrelated children’s weight statuses).19 Univariate analyses of parental ability to correctly identify the weight status of their own child by selected factors were performed using logistic regression. The effect of the relationship between child weight and child age on parental ability to correctly identify the weight status of their own children was also evaluated because of data from previous studies.9,13

Multivariate logistic regression analysis was then applied to identify predictors of parental ability to correctly identify their child’s weight status that are independently significant after adjusting for other variables (Table 5). Receiver operating characteristic evaluation for the logistic regression model was performed.

Statistical analyses were performed on questionnaire responses by using JMP 5.0 statistical software (SAS Institute, Cary, NC) and the R statistical package 4.13.20,21 Significance for all of the analyses was set at P < .05.

RESULTS
A total of 1098 parents were surveyed. Demographics are presented in Table 2. Of those reporting Hispanic ethnicity, 96% were white, 1% were black, and 3% were Asian/other.

Surveyed parents were able to correctly identify the weight status of 7 (6–8) (median [interquartile range]) of 12 presented photographs and of 3 (3–4) of 7 presented photographs of AROW or overweight photographed children. Parents who were women (P = .004) and had >1 child (P = .03) were more likely to correctly assess the weight status of unrelated children compared with category counterparts. Parents were more likely to correctly

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Demographic Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent sex, male:female</td>
<td>141:957</td>
</tr>
<tr>
<td>Parent age, mean ± SD, y</td>
<td>37 ± 8</td>
</tr>
<tr>
<td>Parent BMI, mean ± SD, kg/m²</td>
<td>29.0 ± 7.5</td>
</tr>
<tr>
<td>Parent weight status, %</td>
<td>Obese 36</td>
</tr>
<tr>
<td></td>
<td>Overweight 34</td>
</tr>
<tr>
<td></td>
<td>Normal 29</td>
</tr>
<tr>
<td></td>
<td>Underweight 1</td>
</tr>
<tr>
<td>Parent education status, %</td>
<td>Grade school or less 13</td>
</tr>
<tr>
<td></td>
<td>Middle school 11</td>
</tr>
<tr>
<td></td>
<td>Any high school or high school graduate 20</td>
</tr>
<tr>
<td></td>
<td>Any college education or more 56</td>
</tr>
<tr>
<td>Child sex, male:female</td>
<td>495:603</td>
</tr>
<tr>
<td>Child age, mean ± SD, y</td>
<td>8 ± 4</td>
</tr>
<tr>
<td>Child weight status, %</td>
<td>Overweight 21</td>
</tr>
<tr>
<td></td>
<td>AROW 18</td>
</tr>
<tr>
<td></td>
<td>Normal weight 59</td>
</tr>
<tr>
<td></td>
<td>Underweight 2</td>
</tr>
<tr>
<td>Race, %</td>
<td>White 74</td>
</tr>
<tr>
<td></td>
<td>Asian 7</td>
</tr>
<tr>
<td></td>
<td>Black 12</td>
</tr>
<tr>
<td></td>
<td>Other 7</td>
</tr>
<tr>
<td>Hispanic ethnicity, %</td>
<td>46</td>
</tr>
<tr>
<td>No. of children in family, %</td>
<td>1 18</td>
</tr>
<tr>
<td></td>
<td>2 41</td>
</tr>
<tr>
<td></td>
<td>3 25</td>
</tr>
<tr>
<td></td>
<td>≥4 16</td>
</tr>
<tr>
<td>Income, % above poverty</td>
<td>67</td>
</tr>
<tr>
<td>Home language, %</td>
<td>English only 42</td>
</tr>
<tr>
<td></td>
<td>Spanish only 29</td>
</tr>
<tr>
<td></td>
<td>Bilingual 24</td>
</tr>
<tr>
<td></td>
<td>Other 5</td>
</tr>
<tr>
<td>No. of family generations in the United States, %</td>
<td>1 45</td>
</tr>
<tr>
<td></td>
<td>2 13</td>
</tr>
<tr>
<td></td>
<td>3 9</td>
</tr>
<tr>
<td></td>
<td>≥4 33</td>
</tr>
</tbody>
</table>
assess the weight status of unrelated children if the photographed children were ≥4 years of age but were less likely to correctly assess the weight status of unrelated children if they were male, Hispanic, or AROW or overweight (Table 3). In contrast, race of photographed children did not significantly affect the ability of surveyed parents to correctly assess weight status.

Sixty-one percent of surveyed parents correctly identified their own child’s weight status. Among parents of AROW or overweight children, 30% correctly identified their own child’s weight status. Parents who were non-Hispanic, of normal weight or underweight, had normal weight or underweight children ≥4 years of age, spoke English at home, had received more than a high school education, with incomes above poverty criteria, and whose families had been in the United States for ≥2 generations were more likely than comparison parents to correctly identify the weight status of their own children (P < .05; Table 4).

Parental ability to correctly assess the weight status of their children was not related to their ability to correctly assess the weight status of unrelated children (P = .71). Selection of normal ideal body types among various aged-unrelated children by parents was also not related to parental ability to correctly identify the weight status of their own children (P > .15 for all age categories; Table 4).

Multivariate Regression Analysis
Multiple logistic regression analysis of parental assessment of their own child’s weight status identified significant associations between parental assessment of their child’s weight status and child characteristics. Results are displayed in Table 5. Area under the curve from the receiver operating characteristic was 0.78 for the model.

DISCUSSION
To our knowledge, we present the first study evaluating parents’ perceptions of weight of unrelated children and their relationship with parents’ perceptions of weight of their own children. Our study also assessed a number of socioeconomic and race/ethnic variables to control for potential confounding factors. Although we did demonstrate that similar child variables (i.e., child age and actual weight status) may have affected parental ability to determine the weight status of both related and unrelated children, we did not find a direct association between parental assessment of their children’s weight status and parental assessment of the weight status of unrelated children. The implications of our findings are discussed below.

In our cohort, parental assessments of the weight of their children were associated with primarily child characteristics. In particular, correct parental assessment of the weight of their child was associated with child age and weight status. Parents of children who were school aged-unrelated children by parents was also not related to parental ability to correctly identify the weight status of their own children (P > .05; Table 4).

**TABLE 3** Odds Ratio for Parental Ability to Correctly (Versus Incorrectly) Identify the Weight Status of Unrelated Children According to Photographed Child Variables

<table>
<thead>
<tr>
<th>Photographed Child Characteristics</th>
<th>Unadjusted OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (infant or toddler)</td>
<td>1.63 (1.54–1.73)</td>
</tr>
<tr>
<td>Race (white)</td>
<td>1.00 (0.94–1.08)</td>
</tr>
<tr>
<td>Ethnicity (non-Hispanic)</td>
<td>0.90 (0.84–0.96)</td>
</tr>
<tr>
<td>Sex (female)</td>
<td>0.48 (0.45–0.51)</td>
</tr>
<tr>
<td>Weight status (normal weight or underweight)</td>
<td>0.35 (0.33–0.38)</td>
</tr>
</tbody>
</table>

The reference population is presented in parentheses. OR indicates odds ratio; CI, confidence interval.

**TABLE 4** Odds Ratios for Parental Ability to Correctly (Versus Incorrectly) Identify the Weight Status of Their Own Children, According to Selected Variables

<table>
<thead>
<tr>
<th>Category</th>
<th>Unadjusted OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race (white)</td>
<td>0.87 (0.66–1.14)</td>
</tr>
<tr>
<td>Ethnicity (non-Hispanic)</td>
<td>0.69 (0.55–0.88)</td>
</tr>
<tr>
<td>Parent sex (female)</td>
<td>1.19 (0.83–1.73)</td>
</tr>
<tr>
<td>Parent age (&lt;30 y)</td>
<td>1.16 (0.85–1.56)</td>
</tr>
<tr>
<td>Parent weight status (normal weight or underweight)</td>
<td>0.60 (0.45–0.78)</td>
</tr>
<tr>
<td>Child sex (female)</td>
<td>1.00 (0.78–1.28)</td>
</tr>
<tr>
<td>Child age (&lt;4 y)</td>
<td>1.46 (1.09–1.95)</td>
</tr>
<tr>
<td>Child weight status (normal weight or underweight)</td>
<td>0.10 (0.07–0.13)</td>
</tr>
<tr>
<td>Child age (&lt;4 y) × child weight status (normal weight or underweight)</td>
<td>3.15 (1.81–6.59)</td>
</tr>
</tbody>
</table>

The reference population is presented in parentheses. OR indicates odds ratio; CI, confidence interval.

**TABLE 5** Logistic Regression Analysis of Parental Ability to Correctly Identify Their Child’s Weight Status and Selected Characteristics in the Entire Cohort

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child weight status (normal weight or underweight)</td>
<td>0.04 (0.02–0.07)</td>
</tr>
<tr>
<td>Child age (&lt;4 y)</td>
<td>4.18 (2.39–7.83)</td>
</tr>
<tr>
<td>Child weight status (normal weight or underweight) × child age (&lt;4 y)</td>
<td>3.15 (1.81–6.59)</td>
</tr>
</tbody>
</table>

The reference population is presented in parentheses. OR indicates odds ratio; CI, confidence interval.

a P < .01. b P < .05.

* Results were adjusted for child age and child weight status.
aged (≥4 years) and of normal or underweight status were more likely to correctly assess the weight status of their children as compared with parents of younger and AROW or overweight children. Independent investigations have also demonstrated similar associations between child age and weight status and parental assessment of their child’s weight.7-13,15 Several investigators10,12,15 have also reported that there may be differences in how mothers interpret the weight status of their children according to their children’s sex. Baughcum et al8 demonstrated a significant association between education level and preschool mothers’ assessment of the weight status of their children. We did not demonstrate any association between parents’ assessments of weight of their children and child sex. Although we did demonstrate univariate associations between parents’ perceptions of weight of their children and parent education level, this association was not found to be significant once other socioeconomic demographics were entered into a multivariate model. Possible explanations of the discrepancies between our results and the outcomes of other studies might include our evaluation of a much broader age range of children, inclusion of fathers, and our cohort size.

In our cohort, parental ability to assess the weight status of unrelated children was associated with child characteristics, including age, ethnicity, sex, and weight status. Specifically, parents were more likely to correctly assess the weight status of unrelated children who were ≥4 years of age, non-Hispanic, female, and normal or underweight as compared with their counterparts. Parental ability to correctly identify the weight status of unrelated children may have reflected their exposure to children with these characteristics as reflected by the majority of child participants being ≥4 years of age, non-Hispanic, female, and normal or underweight. In addition, we demonstrated that mothers and parents of families with ≥2 children were more likely than fathers and parents of only 1 child to correctly identify the weight status of photographed children. Assuming that mothers were the primary caregivers to the children represented in this study, as reflected by their attendance at scheduled pediatric care visits, we propose that the comparatively better ability of mothers and parents of ≥2 children to discern the weight status of unrelated children found in our cohort may reflect increased exposure of these individuals to children of different weights and ages.

In our study, parents were more likely to correctly identify the weight status of school-aged (≥4 years old) and normal and underweight children than of younger, AROW, or overweight children, independent of their affiliation with the children. This finding, taken in conjunction with previous studies demonstrating a reduced ability for parents to identify the overweight status of younger children,9,13 suggests a general public misperception of what constitutes normal weight for younger children. Recent studies demonstrate a growing prevalence of overweight and AROW among toddlers and preschool children.22,23 In addition, a review of longitudinal, prospective studies evaluating the relationship between obesity in childhood and adulthood demonstrates that the risk of adult obesity is increased among overweight infants and toddlers.24 Our results and similar findings by other investigators identify a need for public reorientation regarding the definition of appropriate weight for infants and toddlers and suggest that early intervention during the first 2 years of life should be incorporated into prevention and treatment campaigns for childhood obesity.

Although our study data demonstrated that child age and weight status similarly affect parental ability to assess the weight status of unrelated and related children, we also demonstrated that parental assessment of the weight status of their children was not related to parental assessment of the weight status of unrelated children or to parents’ ideal body type selections among unrelated children. Our findings suggest that although parents may concur with public weight perceptions for younger and overweight children, they may also demonstrate weight perceptions unique and specific to their child. Alternatively, the demonstrated lack of application of subjective weight criteria to their own children (we showed no association between parental ability to correctly identify their own child’s weight status and parent-selected body type ideals among unrelated children) may reflect parents’ frank denial of or unwillingness to accept their children’s weight status. Parents may also use different criteria to define overweight status or to assign health-related concerns for their children. Jain et al14 demonstrated that mothers did not believe that their children were overweight if their children were active and had a healthy diet and/or a good appetite. Also, a recent study by Eckstein et al9 found that parental level of health-related concern for their child was more associated with their perception of the child’s activity level than the child’s weight status. Taken together, individualized interventions may be required to help parents uniformly recognize weight status, and particularly overweight, in all of their children.

Parental involvement is important for weight reduction and healthy weight maintenance in children. Parents have a strong influence on children’s dietary intake and level of activity, and parental encouragement has been shown to be important for adoption of healthy eating and physical activity behaviors by their children.25,26 Although not all overweight treatment studies have found that parental participation significantly improves treatment results, the interventions documenting the largest and longest-term decreases in percentage of overweight include parental participation as an integral component.27,28 Crucial to parental involvement in
weight maintenance efforts among children is parental recognition of overweight in their children and a height-
ened level of health concern for their overweight chil-
dren. Studies demonstrate that parental recognition of 
the overweight status of their children and of health risk 
associated with overweight is associated with parental 
readiness for action with regard to their child’s weight.29 
Interventions targeting parental recognition of their 
child’s weight status may, thus, be important for healthy 
weight maintenance in children.

Pediatricians and other clinicians who evaluate fam-
ilies on a regular basis may be ideally situated to promote 
and encourage healthy weight perceptions and parental 
awareness of their child’s weight status given their per-
sonal interaction with parents and children on a recur-
rent basis. The frequent well-child visits typically sched-
uled in the first 2 years of life between pediatricians and 
families may provide the early opportunities needed for 
preventative and treatment interventions among chil-
dren at risk. Physician involvement has been shown to 
promote weight loss behaviors and efforts30–33 and 
weight loss success among adult patients.31 Published 
expert panel guidelines currently advocate for screening 
for childhood overweight via the clinician’s office and 
regular clinician-family discussions regarding weight 
management.17 However, recent data suggest that phy-
sician adoption of expert recommendations for the treat-
ment and prevention of childhood overweight are still 
suboptimal.34,35 Therefore, additional research is needed to 
understand how best to implement early screening 
practices for overweight among children and encourage 
clinician discussions with families regarding what is an 
appropriate and healthy weight for their child.

LIMITATIONS
The findings of this study are subject to a number of 
limitations. First, limited representation of body types in 
our photograph questionnaire (only 3 photographs per 
age group) may not have adequately represented paren-
tal ability to assess the weight status of unrelated chil-
dren or body type ideals in each age category. Although 
we restricted the number of photographs presented to 
reduce the time requirement of the study for feasibility, 
we nevertheless did represent a wide variety of body 
types in our questionnaire, as well as a broad spectrum of 
child characteristics (ie, by race, ethnicity, and sex). In 
addition, we chose to represent the normal weight body 
type in our photograph questionnaire (only 1 per age 
group) by children who plotted at the midrange of nor-
mal. Thus, selection of children in other weight catego-
ries (eg, AROW or overweight) in each age group would 
have represented a true deviation from the norms for 
age. Furthermore, whereas other studies have used fig-
ure sketches without exact correlation to actual BMI-
for-age and sex percentiles to represent body types for 
selection by survey participants, we provided actual rep-
resentations (photographs) of children with known 
BMI-for-age and sex percentiles for selection. Thus, we 
were able to accurately compare and interpret demon-
strated differences in perceived and actual weight status. 
Second, our analyses only accounted for parental ability 
to recognize the weight status of 1 of their children. 
Therefore, for parents with >1 child, we were not able to 
comment on whether ability to recognize weight status 
differed between children or whether assessment of 1 
child was representative of their ability to assess the 
weight status of all of their children. Third, our study 
represents a cross-sectional evaluation and therefore 
cannot comment on whether current parent ability to 
recognize the weight status of their children at a given 
age adequately reflects their ability to recognize the 
weight status of their children in the past or future. 
Additional study is therefore needed to address these 
remaining concerns.

CONCLUSIONS
We demonstrate that parental assessment of their child’s 
weight status is not associated with parental ability to 
determine the weight status of unrelated children. Par-
ents’ perceptions of children’s weight are influenced by 
child characteristics. Pediatricians and other clinicians 
who regularly evaluate children should monitor and 
discuss the weight status of children with their parents as 
early as possible to promote parental awareness of the 
weight status of their children.

ACKNOWLEDGMENTS
This research was supported by a grant from the Rest-
haven Children’s Health Fund and also partially sup-
ported by General Clinical Research Center grant NIH 
M01 RR000827.

We gratefully acknowledge our research volunteers 
and referring medical providers of the Children’s Pri-
mary Care Medical Group (especially Marvin Zaguli, 
MD, John Hansen, MD, Marshall Littman, MD, and 
Robert Bjork, MD).

REFERENCES
1. Ogden CL, Flegal KM, Carroll MD, Johnson CL. Prevalence and 
trends in overweight among US children and adolescents, 
glucose tolerance among children and adolescents with marked 
1989;115:892–897
Lavine JE. Obesity, insulin resistance, and other clinicopatho-
logical correlates of pediatric nonalcoholic fatty liver disease. J 
Pediatr. 2003;143:500–505
5. Sorof J, Daniels S. Obesity hypertension in children: a problem 
6. Schwimmer JB, Burwinkle TM, Varni JW. Health-related qual-
34. O’Brien SH, Holubkov R, Reis E. Identification, evaluation, and management of obesity in an academic primary care center. *Pediatrics.* 2004;114(2). Available at: www.pediatrics.org/cgi/content/full/114/2/e154
Hematologic and Blood Biochemistry Monitoring During Methylphenidate Treatment in Children With Attention-Deficit/Hyperactivity Disorder: 2-Year, Open-Label Study Results

Sharon B. Wigal, PhD*, Timothy E. Wilens, MD*, Mark Wolraich, MD, Marc Lerner, MD*

*Department of Pediatrics, University of California, Irvine, California;Psychopharmacology Research Unit, Massachusetts General Hospital, Boston, Massachusetts;Department of Pediatrics, University of Oklahoma, Oklahoma City, Oklahoma.

ABSTRACT

OBJECTIVES. Patients receiving methylphenidate for the management of attention-deficit/hyperactivity disorder are recommended to receive periodic hematologic monitoring. The objective of this study was to evaluate the long-term effects of methylphenidate treatment on blood biochemistry and hematologic values.

METHODS. This study involved a detailed analysis of changes in hematologic and blood biochemistry values over the course of a 2-year study of once-daily OROS methylphenidate in otherwise healthy children aged 6 to 13 years with attention-deficit/hyperactivity disorder. Routine hematologic and blood biochemistry assessments were performed at baseline, at 6 and 12 months during study treatment, and at the end of the study.

RESULTS. Of the 407 subjects enrolled in the study, 289 completed year 1, and 229 completed 21 of 24 months. No subject was excluded from entry into the study or discontinued from the study because of abnormalities of any of the blood chemistries evaluated. There were no clinically significant changes from baseline in mean values for hematologic or blood biochemistry parameters. For most values, the mean change in value over the course of the study was <5%.

CONCLUSIONS. These longer-term data suggest that chronic therapy with OROS methylphenidate has no clinically significant impact on laboratory values, challenging the necessity of routine hematologic monitoring in otherwise healthy children with attention-deficit/hyperactivity disorder who are treated with methylphenidate.
ATTENTION-DEFICIT/HYPERACTIVITY DISORDER (ADHD) is the most common neurobehavioral disorder treated in children. The disorder is often chronic, with symptoms and associated impairment persisting into adolescence and adulthood. Throughout the life span, ADHD is associated with significant functional impairment, including school failure; comorbid psychiatric and developmental conditions; and family, peer, and emotional difficulties.

Because of the chronic nature of the disorder, long-term treatment has been recommended in ADHD treatment guidelines. A 14-month, long-term, multisite, multimodal treatment study demonstrated that medication management of ADHD is the most important variable in successful treatment outcome. Because of their established efficacy, stimulants are considered first-line agents for the treatment of ADHD. Methylphenidate is the most commonly prescribed stimulant used to treat ADHD, with extended-release formulations, such as OROS methylphenidate (Concerta, McNeil Pediatrics, Ft Washington, PA), providing ADHD symptom management throughout the day with 1 daily dose.

Treatment guidelines recommend routine monitoring for children treated with stimulant medications. Along with the regular assessments of treatment efficacy and the emergence of adverse effects, it has also been recommended that physicians routinely monitor height, weight, blood pressure, and heart rate in patients treated with stimulants. In addition, routine hematologic monitoring, including periodic complete blood cell count, differential, and platelet counts, is advised during prolonged therapy with both immediate-release and extended-release methylphenidate formulations. Rare cases of thrombocytopenia and/or easy bruising, epistaxis, and gingival bleeding; leukopenia; anemia; and eosinophilia have been reported in patients receiving methylphenidate, but a causal relationship with the drug has not been established.

Despite the recommendations of routine hematologic monitoring in methylphenidate-treated patients, over years of extensive methylphenidate use, no known serious blood biochemistry or hematologic abnormalities have been reported in clinical trials, including those in extended-release methylphenidate formulations. Therefore, the clinical usefulness of routine hematologic or blood biochemistry monitoring is uncertain. Given the risks and discomfort associated with routine hematologic and blood biochemistry assessments, a systematic evaluation of a large sample of children prospectively followed for the development of abnormal laboratory values associated with long-term methylphenidate use is warranted. Such data would provide useful safety information to assess the use of routine blood testing in patients with ADHD treated with methylphenidate.

The current study includes an analysis of blood biochemistry values, including hematology, measured during the course of a long-term, open-label trial of OROS methylphenidate in which children with ADHD received therapy for ≤24 months. Based on the literature, no clinically significant effects on blood biochemistries in general, and on hematologic findings in particular, were hypothesized to occur in this otherwise healthy group of children and young adolescents. This study provides, to our knowledge, the largest data set evaluating the long-term effect of methylphenidate on producing clinically important effects on hematologic and blood biochemistry values.

METHODS

Subjects and Study Design

Subjects from 14 sites were enrolled in this open-label study for ≤24 months. The full details of the open-label study are reported elsewhere. The primary study assessed the long-term safety and efficacy of OROS methylphenidate in children with ADHD. Multiple measures of ADHD symptoms, vital signs, weight, height, and laboratory results were assessed throughout the 2-year study period. The study involved children, aged 6 to 13 years, who had participated in previous OROS methylphenidate trials. The previous studies included only children who had been receiving methylphenidate either with a positive response or without having experienced a significant adverse event based on parent or physician reports.

The study included otherwise healthy children with ADHD. Therefore, subjects with clinically significant gastrointestinal problems, a history of clinically important electrocardiogram abnormalities or blood pressure measurements, or a coexisting medical condition likely to interfere with the safe administration of methylphenidate as determined by the investigator were excluded. Patients were also excluded if they were currently taking the following medications: sedative/hypnotic agents; anticonvulsant agents; desmopressin; antihistamines containing sedatives; lithium; oral corticosteroids; or monoamine oxidase inhibitors, tricyclic antidepressants, theophylline, or warfarin. Patients were not allowed to take additional doses of methylphenidate or other medications for treatment of ADHD beyond those prescribed in the study.

Although children with psychiatric comorbidities were eligible for inclusion, those with Tourette’s syndrome or family history of Tourette’s syndrome; an ongoing seizure disorder; bipolar disorder; psychotic disorder; marked anxiety, tension, or agitation; a mood or anxiety disorder requiring drug therapy; drug or alcohol abuse within the 6 months before study entry; or an eating disorder were excluded from study entry. Adolescents with a history of nonresponse to methylphenidate or a hypersensitivity or significant intolerance to methylphenidate also were excluded.

Subjects and Study Design

Subjects from 14 sites were enrolled in this open-label study for ≤24 months. The full details of the open-label study are reported elsewhere. The primary study assessed the long-term safety and efficacy of OROS methylphenidate in children with ADHD. Multiple measures of ADHD symptoms, vital signs, weight, height, and laboratory results were assessed throughout the 2-year study period. The study involved children, aged 6 to 13 years, who had participated in previous OROS methylphenidate trials. The previous studies included only children who had been receiving methylphenidate either with a positive response or without having experienced a significant adverse event based on parent or physician reports.

The study included otherwise healthy children with ADHD. Therefore, subjects with clinically significant gastrointestinal problems, a history of clinically important electrocardiogram abnormalities or blood pressure measurements, or a coexisting medical condition likely to interfere with the safe administration of methylphenidate as determined by the investigator were excluded. Patients were also excluded if they were currently taking the following medications: sedative/hypnotic agents; anticonvulsant agents; desmopressin; antihistamines containing sedatives; lithium; oral corticosteroids; or monoamine oxidase inhibitors, tricyclic antidepressants, theophylline, or warfarin. Patients were not allowed to take additional doses of methylphenidate or other medications for treatment of ADHD beyond those prescribed in the study.

Although children with psychiatric comorbidities were eligible for inclusion, those with Tourette’s syndrome or family history of Tourette’s syndrome; an ongoing seizure disorder; bipolar disorder; psychotic disorder; marked anxiety, tension, or agitation; a mood or anxiety disorder requiring drug therapy; drug or alcohol abuse within the 6 months before study entry; or an eating disorder were excluded from study entry. Adolescents with a history of nonresponse to methylphenidate or a hypersensitivity or significant intolerance to methylphenidate also were excluded.
Subjects were initially assigned to 1 of 3 dosing levels of OROS methylphenidate (18, 36, or 54 mg once daily) based on either their dose in the previous study or their methylphenidate dose received before entry to 1 of the OROS methylphenidate studies. OROS methylphenidate dosage could be adjusted upward or downward in 18-mg increments throughout the study to a maximum of 54 mg per day if the investigator deemed it necessary. Subjects could follow their usual dosing pattern, including having their doses reduced, stopping medication on weekends, or taking planned drug holidays.

The prospective trial was terminated by the sponsor with Food and Drug Administration acknowledgment between 21 and 24 months of subject participation for administrative reasons that were unrelated to safety or effectiveness. Because data were available for only a minority of subjects at 24 months, the final end point is referred to as 21/24 months.

Each subject’s parent(s) or guardian(s) was required to give signed informed consent, and subjects aged ≥7 years were also required to give written assent if they were able to write their name. The consent and assent forms, study protocol, and any advertisements for subjects were reviewed and approved by the institutional review board of participating centers before initiation of the study.

Laboratory Tests
Blood samples were collected from patients before medication exposure, during medication exposure, and at the completion of the study. Because the current study included children who had participated in a previous OROS methylphenidate trial, if <4 weeks had elapsed since completing a previous study, laboratory tests were not performed at baseline. At the 6- and 12-month visits, blood was drawn for hematology and chemistry. Laboratory tests also were performed at the final study visit of the 2-year study. Subjects who withdrew from the study prematurely had their hematologic and blood biochemistry values measured as part of the exit criteria, with laboratory tests to be completed within 2 weeks of study termination.

Patients were not instructed to fast before blood samples were collected. Collection times were random within and across subjects, with most laboratory tests being administered in the late afternoon.

A central clinical laboratory (Covance Inc, Indianapolis, IN) was used to analyze blood samples for determination of the laboratory indices. All of the blood samples were collected, centrifuged, and immediately stored in a temperature-monitored storage device at −20°C. All of the samples were shipped at −20°C to a central laboratory.

The standard hematologic laboratory tests included hemoglobin, hematocrit, red blood cell count, white blood cell count with differential, and platelet count. Blood biochemistry analyses included sodium, potassium, chloride, blood urea nitrogen, creatinine, calcium, phosphorus, glucose, total protein, albumin, bilirubin, alkaline phosphatase, alanine aminotransferase (ALT), aspartate aminotransferase (AST), γ-glutamyltransferase (GGT), and lactate dehydrogenase.

At the central clinical laboratory, flow cytometry assays were used to measure red blood cell count and platelets, and manual microscopy was used to measure red blood cell count morphology. Calculation was used to measure hematocrit. A peroxidase assay was used to measure white blood cell count, neutrophils, lymphocytes, monocytes, and eosinophils. Basophil/lobularity was used to measure basophils. Colorimetric assays were used to measure hemoglobin, albumin, alkaline phosphatase, bilirubin, calcium, creatinine, GGT, phosphorus, and protein. Enzymatic assays were used to measure ALT, AST, glucose, lactate dehydrogenase, and blood urea nitrogen. Ion-selective electrode assays were used to measure electrolytes.

Reference values were defined and standard for all testing, and laboratory values that were out of range were identified and repeated at the investigator’s discretion. For this analysis, postbaseline laboratory values were assigned as “clinically significant change” (above or below the reference limits for that laboratory test) or “no clinically significant change” (within reference limits for that laboratory test). All of the clinically important out-of-range laboratory values were followed until they returned within the reference range, with the investigator treating the patient as medically required at appropriate intervals until such laboratory values stabilized. Patients were to be discontinued from the study if the investigator deemed any clinically significant abnormal laboratory values as possibly attributable to the use of methylphenidate.

Analyses
All of the analyses were completed by using the intent-to-treat patient population data from the last observation carried forward. For each parameter, mean and SD values and change from baseline were determined at each time point. In addition, the percentages of patients with low, normal, or high values at each time point were determined. Reference values were standard for all of the sites and followed by the central laboratory, and values outside of these were defined as high or low.

McNemar’s Test for Marginal Homogeneity
In epidemiologic studies, outcomes are frequently analyzed as falling above or below a fixed value, but with paired data and a fixed value, McNemar’s test of marginal homogeneity may be used to test for equality of the distribution of the outcomes. To complete an additional
analysis of these study data, the percentage of subjects with clinically significant changes in each blood biochemistry parameter at any time point during the study was determined. McNemar’s test for marginal homogeneity was used to determine whether the probability of having an abnormal laboratory test value was equal at baseline and at the end of the study.

Safety

Throughout the study, parents were questioned about adverse events at each follow-up visit. No systematic scales were used for the determination of adverse events. The investigator determined whether any clinically significant out-of-range laboratory values that emerged during treatment with OROS methylphenidate were considered to be adverse events.

RESULTS

Patients

A total of 407 children entered the study, and 289 completed the first year of therapy. Of these, 278 enrolled in the second year of treatment, and 229 children received treatment to the 21/24-month end point. As reported previously, select baseline characteristics and demographic data for all of the enrolled subjects are reported previously, select baseline characteristics and demographic data for all of the enrolled subjects are provided in Table 1.

Dosing

At the start of the study, 47.4% of the subjects were receiving a 36-mg dose of OROS methylphenidate, 28.5% were receiving an 18-mg dose, and 24.1% were receiving a 54-mg dose. Throughout the study, the mean daily dose of OROS methylphenidate increased by 26%, from 35.2 mg at baseline to 44.2 mg at end point (month 21/24). The mean duration of exposure to study medication at any dose was 471 days. As a result of planned medication breaks being permitted during the 2-year study, 163 subjects (40%) reported a period of ≥7 days without medication (90 subjects reported medication breaks of 7–29 days in duration, and 73 subjects reported breaks of >29 days).

Changes From Baseline in Hematologic and Blood Biochemistry Values

At baseline, 98% of hematologic and blood chemistry values were available, and 88% of these values were available at completion of the study. Mean baseline values for parameters measured and mean change from baseline at the final time point of the study are given in Table 2. There were no clinically significant changes from baseline in mean values for hematologic or blood chemistry parameters.

The percentage change from baseline was <5% for most values and only exceeded 5% for monocytes (change from baseline: −0.57% [9.2%]), basophils (change from baseline: −0.16% [21.3%]), eosinophils (change from baseline: −0.36% [11.6%]), hematocrit (change from baseline: 1.90 mL/dL [5.2%]), serum creatinine (change from baseline: 0.09 mg/dL [18.8%]), total bilirubin (change from baseline: −0.1 mg/dL [21.7%]), serum lactate dehydrogenase (change from baseline: −14.67 U/L [6.7%]), and serum alkaline phosphatase (change from baseline: 26.02 U/L [11.0%]). For most parameters, including all of those where the percentage change from baseline exceeded 5%, the absolute change from baseline was less than or approximately the same as the SD for baseline values.

Changes in hematologic and blood biochemistry values were also analyzed as to whether a clinically significant change occurred at any point during the study. For most parameters, the percentage of patients having clinically significant changes in any value at any time during the study was <10%. The percentage of patients was >10% for 6 values: serum glucose (21.4%; n = 65), hematocrit (20.7%; n = 59), total bilirubin (19.5%; n = 64), neutrophils (19.0%; n = 55), lymphocytes (17.0%; n = 49), and white blood cell count (11.1%; n = 35).

Hematologic and blood biochemistry values also were analyzed according to whether they were lower than the reference range (low), within the reference range (normal), or higher than the reference range (high) at baseline and at the final study time point (Table 3). For most parameters, values remained normal or normalized in ≥95% of patients. The only parameters for which >5% of subjects had changes to values outside of the reference range at end of study were bilirubin (19%; n = 39), hematocrit (14%; n = 28), lymphocytes (8%; n = 17), neutrophils (8%; n = 17), serum glucose (8%; n = 16), serum calcium (7%; n = 15), albumin (7%; n = 14), and serum alkaline phosphatase (6%; n = 13).

Approximately 7% of laboratory study visits were recorded as “retest” or “unscheduled visit.” It is not certain whether any or all of these visits were because of investigators requesting follow-up for an out-of-range laboratory value.

TABLE 1 Baseline Characteristics and Demographics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All Treated Subjects (n = 407)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD), y</td>
<td>9.2 (1.8)</td>
</tr>
<tr>
<td>Gender, % male</td>
<td>83</td>
</tr>
<tr>
<td>Race, % White</td>
<td>86</td>
</tr>
<tr>
<td>Black</td>
<td>6</td>
</tr>
<tr>
<td>Asian</td>
<td>1</td>
</tr>
<tr>
<td>Hispanic</td>
<td>4</td>
</tr>
<tr>
<td>Other</td>
<td>3</td>
</tr>
<tr>
<td>Weight, mean (SD), kg</td>
<td>34 (10)</td>
</tr>
<tr>
<td>Height, mean (SD), cm</td>
<td>137 (11)</td>
</tr>
</tbody>
</table>

McNemar's Test for Marginal Homogeneity

Although all of the patients had laboratory values within the reference ranges at baseline, McNemar’s test for marginal homogeneity was used to assess whether the probability of having a value outside reference limits was greater at the last time point in the study compared with at baseline (Table 4). The results indicate that for most parameters, there was no statistically significant change in the probability of abnormal values. The only parameters that yielded statistically significant P values of \( P < 0.05 \) were hematocrit (\( P = 0.042 \)), red blood cell count (\( P = 0.008 \)), serum calcium (\( P = 0.041 \)), serum albumin (\( P = 0.022 \)), and total bilirubin (\( P < 0.001 \)).

Clinical Laboratory Adverse Events

Before enrollment in the 2-year, open-label study, no subjects were excluded from the original controlled clinical trials because of abnormal laboratory findings, and no subjects were dropped from the controlled studies because of clinically significant laboratory findings or adverse events. Likewise, no subjects developed a clinically significant laboratory value during the course of the 2-year open study necessitating discontinuation from the trial.

Over the 2-year period, 3 patients had abnormal hematologic/blood biochemistry laboratory evaluations that were reported as adverse events per the investigators’ discretion. Leukopenia was reported in 1 patient while taking 36 mg of OROS methylphenidate (white blood cell count \( = 4000/\mu L \) at baseline, \( 2790/\mu L \) at 12 months, and \( 3430/\mu L \) at study end point). One patient was reported to have increased creatinine while taking 36 mg of OROS methylphenidate (creatinine \( = 0.6 \text{ mg/dL} \) at baseline, \( 0.5 \text{ mg/dL} \) at 12 months, and \( 1.2 \text{ mg/dL} \) at study end point). Abnormal liver function test results were reported in 1 patient taking 54 mg of OROS methylphenidate (AST \( = 40.0 \text{ U/L} \) at baseline and \( 53.0 \text{ U/L} \) at study end point; ALT \( = 44.0 \text{ U/L} \) at baseline and \( 51.0 \text{ U/L} \) at study end point; GGT \( = 13.0 \text{ U/L} \) at baseline and \( 26.0 \text{ U/L} \) at study end point). In all 3 of the patients, the adverse events were considered mild, continuous, and possibly related to OROS methylphenidate treatment. The subjects with elevated creatinine and liver enzymes were not discontinued from the study and were referred to their primary care physicians. No follow-up was reported by the study investigators, and no other information is available.

DISCUSSION

The results of these analyses were reassuring in that no subject developed a clinically significant laboratory value during the course of the 2-year open study that neces-

### TABLE 2 Change From Baseline in Hematologic and Blood Biochemistry Values

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Baseline Value, Mean (SD)</th>
<th>Change From Baseline, Mean (% Change)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hematology</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White blood cell count, 1000s/µL</td>
<td>7.03 (2.07)</td>
<td>-0.19 (2.7)</td>
</tr>
<tr>
<td>Neutrophils, %</td>
<td>50.17 (9.74)</td>
<td>2.01 (4.0)</td>
</tr>
<tr>
<td>Lymphocytes, %</td>
<td>39.8 (9.20)</td>
<td>-1.03 (2.6)</td>
</tr>
<tr>
<td>Monocytes, %</td>
<td>6.19 (1.78)</td>
<td>-0.57 (9.2)</td>
</tr>
<tr>
<td>Basophils, %</td>
<td>0.75 (0.35)</td>
<td>-0.16 (21.3)</td>
</tr>
<tr>
<td>Eosinophils, %</td>
<td>3.09 (2.87)</td>
<td>-0.36 (11.6)</td>
</tr>
<tr>
<td>Platelet count, 1000s/µL</td>
<td>288.26 (58.77)</td>
<td>-4.89 (1.7)</td>
</tr>
<tr>
<td>Hemoglobin, g/dL</td>
<td>13.0 (0.81)</td>
<td>0.51 (3.9)</td>
</tr>
<tr>
<td>Hematocrit, mL/dL</td>
<td>36.72 (2.47)</td>
<td>1.90 (5.2)</td>
</tr>
<tr>
<td>Red blood cell count, million per µL</td>
<td>4.61 (0.32)</td>
<td>0.12 (2.6)</td>
</tr>
<tr>
<td><strong>Blood biochemistry</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum sodium, mEq/L</td>
<td>138.42 (2.06)</td>
<td>2.51 (1.8)</td>
</tr>
<tr>
<td>Serum potassium, mEq/L</td>
<td>4.08 (0.34)</td>
<td>0.06 (1.5)</td>
</tr>
<tr>
<td>Serum chloride, mEq/L</td>
<td>101.95 (2.25)</td>
<td>2.52 (2.5)</td>
</tr>
<tr>
<td>Serum glucose, mg/dL</td>
<td>91.04 (13.74)</td>
<td>-0.02 (0.2)</td>
</tr>
<tr>
<td>Blood urea nitrogen, mg/dL</td>
<td>13.01 (3.17)</td>
<td>0.37 (2.8)</td>
</tr>
<tr>
<td>Creatinine (rate blanked), mg/dL</td>
<td>0.48 (0.10)</td>
<td>0.09 (18.8)</td>
</tr>
<tr>
<td>Calcium (EDTA), mg/dL</td>
<td>9.59 (0.36)</td>
<td>0.19 (2.0)</td>
</tr>
<tr>
<td>Phosphorus, mg/dL</td>
<td>4.83 (0.46)</td>
<td>0.12 (2.5)</td>
</tr>
<tr>
<td>Total protein, g/dL</td>
<td>7.06 (0.40)</td>
<td>0.22 (3.1)</td>
</tr>
<tr>
<td>Albumin, g/dL</td>
<td>4.26 (0.25)</td>
<td>0.12 (2.8)</td>
</tr>
<tr>
<td>Total bilirubin, mg/dL</td>
<td>0.46 (0.22)</td>
<td>-0.1 (21.7)</td>
</tr>
<tr>
<td>GGT, U/L</td>
<td>11.72 (4.16)</td>
<td>0.31 (2.6)</td>
</tr>
<tr>
<td>Lactate dehydrogenase, U/L</td>
<td>219.78 (33.01)</td>
<td>-14.67 (6.7)</td>
</tr>
<tr>
<td>Alkaline phosphatase, U/L</td>
<td>235.52 (61.09)</td>
<td>26.02 (11.0)</td>
</tr>
<tr>
<td>AST, U/L</td>
<td>27.95 (5.56)</td>
<td>0.79 (2.8)</td>
</tr>
<tr>
<td>ALT, U/L</td>
<td>17.83 (7.31)</td>
<td>0.59 (3.3)</td>
</tr>
</tbody>
</table>

* This was the last laboratory test result during the 2-year study for patients with baseline and end-point data available.
sitated discontinuation from the trial. Moreover, the results of the study indicate that for most hematologic and blood biochemistry parameters, the change in mean value over the course of the study was <5%, and shifts to abnormal values (high or low) were seen in <5% of subjects. These data suggest that routine blood monitoring is unwarranted in otherwise healthy youth with ADHD receiving OROS methylphenidate treatment. The effects of methylphenidate on youth with ADHD with baseline medical conditions (ie, hepatitis, bone marrow suppression, etc) remain unclear.

Our overall findings indicate that most values over the course of the current study were within the reference range, with a small percentage falling outside reference values. This would be expected because reference ranges for laboratory values are set to include 95% of healthy subjects, with 5% of healthy subjects falling outside of the reference range. For most values with subjects within the reference range at baseline, <10% of subjects experienced clinically significant shifts at any time during the study. Of the 5 values that showed a statistically significant change, only hematocrit and total bilirubin showed clinically significant changes in >10% of subjects and shifts to abnormal values over the course of the study in >5% of subjects. For all of the values, the absolute change from baseline to end point was either approximately the same as or less then the SD for the baseline value. These data suggest that although prolonged therapy with OROS methylphenidate may produce statistically significant changes in such laboratory values as hematocrit and bilirubin, it does not produce clinically significant (or important) changes in hematologic or blood biochemistry values overall.

Hematologic and blood biochemistry values differ by age, gender, and race, and such differences are associated with maturation. For example, in the Bogalusa Heart Study, complete blood cell counts were analyzed from 3018 children and adolescents aged 5 to 17 years and demonstrated that hemoglobin levels, hematocrit, and red blood cell count increased with age ($P < .01$), and white blood cell count decreased with age ($P < .0005$). Because a naturalistic fluctuation in values is usually seen in children, especially as they age, such variation would be expected throughout the course of a

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Reference Range</th>
<th>Subjects With Values Within the Reference Range at Baseline, n (%)</th>
<th>Subjects With Normal or Normalized Values at End of Study, n/N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hematology</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White blood cell count, 1000s/μL</td>
<td>4.35–13.65</td>
<td>316 (93.5)</td>
<td>196/206 (95)</td>
</tr>
<tr>
<td>Neutrophils, %</td>
<td>40.5–75.0</td>
<td>290 (85.8)</td>
<td>189/206 (92)</td>
</tr>
<tr>
<td>Lymphocytes, %</td>
<td>15.4–48.5</td>
<td>288 (85.2)</td>
<td>189/206 (92)</td>
</tr>
<tr>
<td>Monocytes, %</td>
<td>2.6–10.1</td>
<td>329 (97.3)</td>
<td>203/206 (99)</td>
</tr>
<tr>
<td>Basophils, %</td>
<td>0–2.0</td>
<td>337 (99.7)</td>
<td>205/206 (99)</td>
</tr>
<tr>
<td>Eosinophils, %</td>
<td>0–6.8</td>
<td>301 (89.1)</td>
<td>196/206 (95)</td>
</tr>
<tr>
<td>Platelet count, 1000s/μL</td>
<td>130–394</td>
<td>326 (96.4)</td>
<td>202/206 (98)</td>
</tr>
<tr>
<td>Hemoglobin, g/dL</td>
<td>12.7–18.1</td>
<td>329 (97.3)</td>
<td>198/206 (96)</td>
</tr>
<tr>
<td>Hematocrit, mL/dL</td>
<td>34–44</td>
<td>285 (87.4)</td>
<td>172/200 (86)</td>
</tr>
<tr>
<td>Red blood cell count, million per μL</td>
<td>4.5–6.4</td>
<td>332 (98.2)</td>
<td>199/206 (97)</td>
</tr>
<tr>
<td>Blood biochemistry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum sodium, mEq/L</td>
<td>132–147</td>
<td>338 (100.0)</td>
<td>204/204 (100)</td>
</tr>
<tr>
<td>Serum potassium, mEq/L</td>
<td>3.4–5.4</td>
<td>328 (99.4)</td>
<td>199/202 (99)</td>
</tr>
<tr>
<td>Serum chloride, mEq/L</td>
<td>94–112</td>
<td>338 (100.0)</td>
<td>204/204 (100)</td>
</tr>
<tr>
<td>Serum glucose, mg/dL</td>
<td>70–115</td>
<td>304 (92.1)</td>
<td>187/203 (92)</td>
</tr>
<tr>
<td>Blood urea nitrogen, mg/dL</td>
<td>4–24</td>
<td>338 (100.0)</td>
<td>206/206 (100)</td>
</tr>
<tr>
<td>Creatinine (rate blanked), mg/dL</td>
<td>0.2–0.6</td>
<td>333 (98.5)</td>
<td>201/206 (98)</td>
</tr>
<tr>
<td>Calcium (EDTA), mg/dL</td>
<td>8.4–10.3</td>
<td>331 (97.9)</td>
<td>190/205 (93)</td>
</tr>
<tr>
<td>Phosphorus mg/dL</td>
<td>2.2–5.1, 3.1–6.0, or 3.2–6.1</td>
<td>330 (100.0)</td>
<td>198/204 (97)</td>
</tr>
<tr>
<td>Total protein, g/dL</td>
<td>6.1–8.4</td>
<td>336 (99.4)</td>
<td>205/205 (100)</td>
</tr>
<tr>
<td>Albumin, g/dL</td>
<td>2.9–4.7</td>
<td>328 (97.3)</td>
<td>191/205 (93)</td>
</tr>
<tr>
<td>Total bilirubin, mg/dL</td>
<td>0.2–1.2</td>
<td>329 (97.6)</td>
<td>167/206 (81)</td>
</tr>
<tr>
<td>GGT, U/L</td>
<td>0–24, 0–35, or 0–51</td>
<td>337 (99.7)</td>
<td>204/205 (99)</td>
</tr>
<tr>
<td>Lactate dehydrogenase, U/L</td>
<td>145–300</td>
<td>319 (96.7)</td>
<td>201/203 (99)</td>
</tr>
<tr>
<td>Alkaline phosphatase, U/L</td>
<td>51–300</td>
<td>322 (95.3)</td>
<td>192/205 (94)</td>
</tr>
<tr>
<td>AST, U/L</td>
<td>10–40</td>
<td>327 (97.0)</td>
<td>199/204 (98)</td>
</tr>
<tr>
<td>ALT, U/L</td>
<td>6–43</td>
<td>331 (98.2)</td>
<td>200/204 (98)</td>
</tr>
</tbody>
</table>

Note: These were the last laboratory test results during the 2-year study for patients with baseline and end-point data available. Reference range for laboratory value varies on the basis of gender and age (full range given).
Over the 24-month treatment period, OROS methylphenidate treatment was safe, with only 3 patients having abnormal hematologic/blood biochemistry laboratory evaluations reported as adverse events. The OROS methylphenidate dosage varied. The patient with leukopenia and the patient with increased creatinine were treated with 36 mg of OROS methylphenidate. The patient with abnormal liver function tests (elevated liver enzymes) was treated with 54 mg of OROS methylphenidate. All 3 of the adverse events were considered mild and continuous by the study investigator, and no subject was withdrawn from the study because of adverse events. Although the subjects with elevated creatinine and liver enzymes were referred to their primary care physicians, no follow-up was reported by the study investigators.

This is the first time that such a detailed analysis of hematologic and blood biochemistry values has been reported for any methylphenidate therapy during long-term treatment in subjects with ADHD; however, several other studies have suggested previously that no consistent changes in such values occur. For example, a 3-week, double-blind, placebo-controlled study of modified-release methylphenidate (Metadate CD) in children with ADHD found no consistent changes in hematologic or blood biochemistry values.\(^9\) In fact, fewer children in the modified-release methylphenidate group (48 of 158 [30%]) had out-of-range laboratory values that were within reference range at baseline compared with the placebo group (63 of 163 [39%]). The lack of consistent changes was also confirmed in a 3-week, open-label study in which blood biochemistry was assessed at baseline and weekly throughout.\(^{24}\) In the study, only subjects with blood biochemistry values within reference ranges were included, and there were no clinically significant changes in any of the hematology or blood biochemistry values throughout the study.

As was demonstrated in the MTA study, frequent clinician-patient contact and monitoring of treatment outcome is an important part of ensuring optimal efficacy.\(^8\) Although periodic monitoring of growth (height and weight) and vital signs (blood pressure and heart rate), as well as adverse events and the efficacy of treatment, seem to be necessary,\(^{6,7,15,16}\) routine hematologic monitoring in methylphenidate-treated patients does not. The absence of clinically significant effects of stimulant therapy with methylphenidate on values of hematologic values or blood biochemistry (as demonstrated for OROS methylphenidate in this long-term study) suggests that routine monitoring of hematologic and blood biochemistry laboratory values in otherwise healthy children may not be necessary. Additional work in medically compromised children is needed to address this question completely.

In clinical practice, routine monitoring of hematologic and blood biochemistry parameters is intended to serve as a screening tool to identify individual patients with an adverse reaction. The current study analyzed changes in hematologic and blood biochemistry parameters as a group means and not as individual changes in the parameters measured. Therefore, additional studies are needed to determine the threshold for clinically important changes and to determine how many patients are likely to have elevations of what magnitude above the determined threshold.

The results of this study need to be tempered against their limitations. Children were screened to be otherwise healthy, and, therefore, this is a selected group that may not be representative of general pediatric and psychiatric practices. However, no youth was excluded because of abnormal values. The generalizability of these data to medically compromised children (ie, those with hepatic, hematologic, or other blood dyscrasias) remains unknown. Exclusion or discontinuation from the clinical trial based on “clinically significant findings” varied by

---

**TABLE 4** McNemar’s Test for Marginal Homogeneity: Baseline Versus Last Time Point

<table>
<thead>
<tr>
<th>Parameter</th>
<th>McNemar’s Statistic</th>
<th>(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hematology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White blood cell count, 1000s/μL</td>
<td>0.290</td>
<td>0.590</td>
</tr>
<tr>
<td>Neutrophils, %</td>
<td>0.581</td>
<td>0.446</td>
</tr>
<tr>
<td>Lymphocytes, %</td>
<td>1.209</td>
<td>0.272</td>
</tr>
<tr>
<td>Monocytes, %</td>
<td>0.600</td>
<td>0.439</td>
</tr>
<tr>
<td>Basophils, %</td>
<td>1.000</td>
<td>0.317</td>
</tr>
<tr>
<td>Eosinophils, %</td>
<td>2.077</td>
<td>0.150</td>
</tr>
<tr>
<td>Platelet count, 1000s/μL</td>
<td>1.923</td>
<td>0.166</td>
</tr>
<tr>
<td>Hemoglobin, g/dL</td>
<td>3.200</td>
<td>0.674</td>
</tr>
<tr>
<td>Hematocrit, mL/dL</td>
<td>4.129</td>
<td>0.042</td>
</tr>
<tr>
<td>Red blood cell count, million per μL</td>
<td>7.143</td>
<td>0.008</td>
</tr>
</tbody>
</table>

**Blood biochemistry**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>McNemar’s Statistic</th>
<th>(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serum sodium, mEq/L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum potassium, mEq/L</td>
<td>1.286</td>
<td>0.257</td>
</tr>
<tr>
<td>Serum chloride, mEq/L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum glucose, mg/dL</td>
<td>0.021</td>
<td>0.884</td>
</tr>
<tr>
<td>Blood urea nitrogen, mg/dL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creatinine (rate blanked), mg/dL</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Calcium (EDTA), mg/dL</td>
<td>4.167</td>
<td>0.041</td>
</tr>
<tr>
<td>Phosphorous, mg/dL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total protein, g/dL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Albumin, g/dL</td>
<td>5.261</td>
<td>0.022</td>
</tr>
<tr>
<td>Total bilirubin, mg/dL</td>
<td>41.143</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GGT, U/L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lactate dehydrogenase, U/L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alkaline phosphatase, U/L</td>
<td>2.130</td>
<td>0.144</td>
</tr>
<tr>
<td>AST, U/L</td>
<td>1.000</td>
<td>0.317</td>
</tr>
<tr>
<td>ALT, U/L</td>
<td>0.091</td>
<td>0.763</td>
</tr>
</tbody>
</table>

\(a\) Less than 2 nonmissing levels for treatment and baseline.

\(b\) Less than 2 nonmissing levels for baseline.
ACKNOWLEDGMENTS
We acknowledge the contributions of the Concerta Study Group and the editorial support of JK Associates, Inc.

REFERENCES


Abnormal Thermoregulatory Responses in Adolescents With Chronic Fatigue Syndrome: Relation to Clinical Symptoms

Vegard Bruun Wyller, MDa,b, Kristin Godang, BSc, Lars Markrid, MD, PhD, Jerome Philip Saul, MD, Erik Thaulow, MD, PhD, Lars Walløe, MD, PhD

Departments of aPediatrics, cEndocrinology, and dMedical Biochemistry, Rikshospitalet-Radiumhospital Medical Center, Oslo, Norway; bDepartment of Physiology, University of Oslo, Oslo, Norway; eDepartment of Pediatrics, Medical University of South Carolina, Charleston, South Carolina

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

ABSTRACT

OBJECTIVES. Chronic fatigue syndrome is a common and disabling disease of unknown etiology. Accumulating evidence indicates dysfunction of the autonomic nervous system. To further explore the pathophysiology of chronic fatigue syndrome, we investigated thermoregulatory responses dependent on catecholaminergic effector systems in adolescent patients with chronic fatigue syndrome.

PATIENTS AND METHODS. A consecutive sample of 15 patients with chronic fatigue syndrome aged 12 to 18 years and a volunteer sample of 57 healthy control subjects of equal gender and age distribution were included. Plasma catecholamines and metanephrines were measured before and after strong cooling of 1 hand. Acral skin blood flow, tympanic temperature, heart rate, and mean blood pressure were measured during moderate cooling of 1 hand. In addition, clinical symptoms indicative of thermoregulatory disturbances were recorded.

RESULTS. Patients with chronic fatigue syndrome reported significantly more shivering, sweating, sudden change of skin color, and feeling unusually warm. At baseline, patients with chronic fatigue syndrome had higher levels of norepinephrine, heart rate, epinephrine, and tympanic temperature than control subjects. During cooling of 1 hand, acral skin blood flow was less reduced, vasoconstrictor events occurred at lower temperatures, and tympanic temperature decreased more in patients with chronic fatigue syndrome compared with control subjects. Catecholamines increased and metanephrines decreased similarly in the 2 groups.

CONCLUSIONS. Adolescent patients with chronic fatigue syndrome have abnormal catecholaminergic-dependent thermoregulatory responses both at rest and during local skin cooling, supporting a hypothesis of sympathetic dysfunction and possibly explaining important clinical symptoms.
**Chronic Fatigue Syndrome (CFS)** is a disabling disease that affects people worldwide. The etiology remains unknown. However, there is accumulating evidence of orthostatic intolerance and other disturbances of cardiovascular regulation in this patient group. Dysfunction of the autonomic nervous system, and particularly the sympathetic branch, has, therefore, been proposed as an important component of the pathophysiology. Notably, the autonomic nervous system also has an essential role in human thermoregulation. However, this regulatory function has hardly been explored in CFS despite frequent complaints of fever-like sensation, night sweats, and subjective temperature sensitivity (feeling too hot or cold).

Human thermoregulation involves complex control mechanisms. Discrete regions of the hypothalamus are sensitive to changes in core body temperature; in addition, these regions receive afferent information from peripheral thermoreceptors. The main effector organs are skeletal muscles, sweat glands, skin arterioles, and skin arteriovenous anastomoses (AVAs). Skeletal muscles are controlled by motor neurons of the somatic nervous system; however, enhanced sympathetic outflow and epinephrine secretion promote shivering.

Sweat glands are innervated by cholinergic sympathetic neurons, which also seem to promote vasodilation in nearby arterioles by local release of nitric oxide and prostaglandins from endothelial cells. In addition, skin arterioles are controlled by vasoconstrictive noradrenergic sympathetic neurons. The AVAs, of which their presence seem to be confined to acral, glabrous skin areas (e.g., the palms and soles), receive noradrenergic sympathetic efferents only; thus, AVA blood flow fluctuations reflect sympathetic nerve activity. Although AVAs are influenced by local factors as well, previous experiments in our laboratory indicate that they are under strong central control and that they play an important part in overall thermoregulation.

The aim of this study was to explore aspects of thermoregulation in adolescent patients with CFS. Because CFS has been linked to sympathetic dysfunction, we specifically focused on catecholaminergic-dependent effector systems and hypothesized to find differences between patients with CFS and healthy control subjects both at rest and during local skin cooling.

**Patients and Methods**

**Subjects**

Patients with CFS aged 12 to 18 years were consecutively recruited from the outpatient clinic at the Department of Pediatrics, Rikshospitalet-Radiumhospitalet Medical Center, serving as a national referral center for children and adolescents with unexplained chronic fatigue. Other disease states that might explain their present symptoms, such as autoimmune, endocrine, neurologic, or psychiatric disorders, were ruled out by a thorough and standardized set of investigations. Different case definitions of CFS exist. This study adhered to all of the main criteria in the definition from the Centers for Disease Control and Prevention (CDC). Specifically, we required ≥6 months of chronic or relapsing fatigue, severely affecting daily activities; the fatigue should not be explained by any concurrent condition, it should be new or definite in onset, it should not be related to ongoing exertions, and it should not be alleviated by rest. In addition, according to the CDC definition, the patients are also required to report ≥4 of 8 specific accompanying symptoms (headache, muscle pain, joint pain, sore throat, tender lymph nodes, impaired memory/concentration, unrefreshing sleep, and postexertional malaise). However, the validity of this last demand has been questioned, both in pediatric and adult patients. Indeed, recent evidence, as reviewed by Cho et al, raises serious concerns about this part of the CDC definition. Therefore, accompanying symptoms were not required in this study.

Healthy control subjects aged 12 to 18 years volunteered from local schools. The recruiting procedure ensured an equal distribution of age and gender among patients and control subjects. A high number of participants in both groups would have yielded best statistical power. However, because control subjects were far easier to recruit than patients with CFS, we aimed at an ~4:1 relation. Subjects having a chronic disease (such as allergy, skin diseases, vascular diseases, or diabetes) or using drugs (including contraceptive pills) on a regular basis were excluded.

One week before the experiments, all of the participants were instructed not to drink beverages containing alcohol or caffeine, not to take any drugs, and not to use tobacco products. On the day of the experiments, they were instructed to have fasted overnight.

Written informed consent was obtained from all of the participants and their parents. The study was approved by the Regional Committee for Ethics in Medical Research.

**Questionnaire**

Items from the Autonomic Symptom Profile, a validated instrument for assessing autonomic dysfunction, was translated into Norwegian by Dr Wyller and slightly modified to fit our particular age group. A couple of items in this questionnaire are specifically directed toward sudomotor and vasomotor function, which, in turn, is related to mechanisms of thermoregulation. The subjects answered by interview.

**Neuroendocrine Responses to Strong Cooling**

The experimental part of this study was undertaken at the Department of Pediatrics, Rikshospitalet-Radiumhospitalet Medical Center and was performed during the
same day in each participant. The first set of experiments began at 8:00 AM and was conducted in a quiet room with normal ambient temperature. The subjects were instructed to apply an ointment containing the local anesthetic lidocaine (eutectic mixture of local anesthetic) on the skin of the antecubital fossa 1 hour in advance. They rested supine for ~15 minutes, and a catheter was then placed in an antecubital vein. After supine rest for another 15 minutes, blood samples were collected on ice-cold Vacutainer tubes containing gluta-thione-ethylene glycol tetraacetic acid and heparin for measurements of plasma catecholamines (norepinephrine, epinephrine, and dopamine) and metanephrines (normetanephrine, metanephrine, and 3-methoxytyramine), respectively. The opposite hand was then immersed in cold water (10°C) for 2 minutes, after which new blood samples were collected immediately. Samples were then centrifuged at 4°C, and plasma was separated for storage at ~80°C until assayed.

Metanephrines were extracted from plasma using solid-phase ion exchange columns (Bond Elute-Accucat; Varian Medical Systems, Palo Alto, CA) and a commercial mobile phase (Chromsystems, München, Germany). Both catecholamines and metanephrines were quantified by high-performance liquid chromatography with a reverse-phase column and glassy carbon electrochemical detector (Agilent Technologies, Colorado Springs, CO). The intraassay and interassay variations were 10.7% and 14.5% for norepinephrine, 23.5% and 10.5% for epinephrine, 7.2% and 6.8% for dopamine, 12.6% and 4.3% for normetanephrine, 11.7% and 3.0% for metanephrine, and 15.3% and 11.2% for 3-methoxytyramine, respectively.

Cardiovascular Responses to Moderate Cooling

The second set of experiment started at 2:00 PM. The participants had been offered a standardized, light meal (2 pieces of bread, 1 glass of juice) 1 hour before but were otherwise not allowed to eat or drink. They were lightly dressed and lay supine in a climatic chamber. The ambient temperature was maintained at ~26°C, assuring that subjects were within their thermoneutral zone. The left hand was immersed in a stirred thermostat-controlled water bath (CB 29-20e; Heto-Holtan, Åbyhøj, Denmark) with a temperature of 35°C, which corresponds with thermoneutrality in water. Thirty minutes were used for acclimatization and another 5 minutes for baseline registration, after which the water temperature was gradually lowered to 19°C over ~50 minutes (Fig 1).

Continuous recordings of acral skin laser flux, which is a measure of acral skin blood flow (ASBF), were obtained by a laser-Doppler instrument (DRT4; Moore Instruments, Milway, Devon, United Kingdom). The probes, which function equally well under water, were firmly attached with adhesive strips to the pulp of the flexor surface of the distal phalanges of the right and left index finger. In this position, the probes mainly detect blood flow in AVAs, as has been shown in earlier experiments in our laboratory. Typanic temperature (TT) was continuously monitored by an electronic probe (D-TM1; Exacon, Roskilde, Denmark) inserted in the outer auditory canal and isolated from the ambient air by a piece of cotton. TT corresponds well with core body temperature. Instantaneous heart rate (HR) was obtained from the R-R interval of the electrocardiogram. Photoplethysmography on the right third finger was used to obtain a noninvasive, continuous recording of arterial blood pressure (2300 Finapres; Ohmeda, Madison, WI). This method correlates satisfactorily with invasive pressure measurements and has also been validated in adolescents and children.

All of the recorded signals, including the temperature of the water bath, were transferred online to a recording computer running a program for real-time data acquisition (developed by Morten Eriksen, Department of Physiology, University of Oslo, Oslo, Norway). Beat-to-beat mean blood pressure (MBP) was calculated by numerical integration of the recorded instantaneous blood pressure. The other physiologic variables were also converted to beat-to-beat records.

Data Analysis

Data were exported to Microsoft Excel (Microsoft, Redmond, WA) for further calculations. For the strong-cooling experiments, Δ cooling (postcooling – precooling) was computed for each variable. For the moderate-cooling experiments, the beat-to-beat records of each variable were used to find the median value (of 100 consecutive heart beats) at 35°C water temperature (precooling), 27°C water temperature (early cooling), and 19°C water temperature (late cooling). Δ cooling 1 (early cooling – precooling) and Δ cooling 2 (late cooling – precooling) was also computed for each variable.

During moderate cooling, ASBF in the cooled hand usually ceases quite suddenly, probably reflecting a coordinated closure of the AVAs (Fig 1). In each subject, this vasoconstrictor event was defined to occur at the water temperature where the median value of ASBF left (computed over 30 consecutive heart beats) was permanently <50% of the median value before cooling.

The statistical analyses were conducted by using SPSS statistical software (SPSS Inc, Chicago, IL). On the basis of inspection of plots, most variables were appraised not to follow a normal distribution, and experimental results are, therefore, expressed as medians with nonparametric 95% confidence intervals (CIs) (Tables 3 and 4). The vasoconstrictor events in the moderate-cooling experiments were subjected to “survival analysis” using a Kaplan-Meier plot (Fig 2). Fisher’s exact test, Wilcoxon-Mann-Whitney’s test, and log rank test (all 2-sided) were used to explore differences between the 2 groups.
FIGURE 1
Original recordings of data from 1 control subject (left) and 1 patient with CFS (right) during the moderate-cooling experiments. AU indicates arbitrary units. The vertical lines indicate initiation of cooling from 35°C to 19°C of the left hand. Artifacts in the records have been manually removed and replaced by values obtained by linear interpolation. Note the well-defined vasoconstrictor event in the control subject (arrow). This phenomenon is accompanied by a slight, sustained increase in TT (lower), probably caused by the cessation of blood flow to cold skin areas and, thus, reduced heat loss. Simultaneously, there is a transient increase in HR (top), which might be because of a transient increase in cardiac sympathetic outflow. In the patient with CFS, there is no well-defined vasoconstrictor event, and TT fell almost linearly toward the end of the experiment.
A $P \leq .05$ was considered statistically significant. To reduce the methodologic problem of multiple comparisons, statistical tests were only performed for the variables precooling and cooling 2 (Tables 3 and 4). Among these variables, the changes in norepinephrine, epinephrine, ASBF left, and TT were considered most central to our research questions.

RESULTS

A total of 15 patients with CFS and 57 healthy control subjects were included in the study (Table 1). The 2 groups were comparable regarding gender, age, weight, and height. All were of white ethnicity, except 1 control subject. Mean duration of fatigue among the patients was 31 months.

The patients with CFS reported significantly more shivering, sweating, and feeling unusually warm (Table 2). They also experienced sudden paleness of skin more often than control subjects, whereas sudden redness of skin was less common in the patient group.

Before strong cooling, patients with CFS had significantly higher levels of norepinephrine ($P \leq .01$) and epinephrine ($P \leq .05$) than control subjects (Table 3). There were no differences in the levels of dopamine and metanephrines (normetanephrine, metanephrine, and 3-methoxytyramine). During strong cooling of the opposite hand, levels of norepinephrine and epinephrine increased in both groups, whereas levels of normetanephrine and metanephrine decreased in both groups. However, all of the variables changed similarly, thus not altering the differences already present at baseline.

Before moderate cooling, patients with CFS had higher HR ($P \leq .01$) and TT ($P \leq .05$) than control subjects (Table 4). There was also a tendency toward lower ASBF in both hands, although this was not significant. During cooling of the left hand, ASBF left decreased significantly less in patients with CFS than in control subjects ($P \leq .01$; Table 4 and Fig 1). Thus, at late cooling, patients with CFS tended to have higher ASBF left than control subjects. Accordingly, the vasoconstrictor events occurred at lower temperatures in patients with CFS than in control subjects ($P \leq .01$; Table 4 and Fig 1). The same tendency, although not significant, was observed for ASBF right. With cooling, TT decreased significantly among patients with CFS but remained stable among control subjects ($P \leq .05$), nullifying the differences present at baseline (Table 4 and Fig 1). HR and MBP remained stable in both groups throughout the experiment.

DISCUSSION

The most important findings of this study are that (1) patients with CFS report several symptoms that might indicate thermoregulatory disturbances; (2) at baseline, patients with CFS have higher levels of norepinephrine and epinephrine and higher TT than control subjects; and (3) during cooling of 1 hand, the neuroendocrine responses are similar in the 2 groups, but ASBF is less reduced among patients with CFS, whereas the baseline differences in TT disappear. Furthermore, the relevance of these findings is strengthened by the strikingly homogeneous responses within the CFS group, creating significant differences from control subjects despite the small number of subjects studied.

Baseline Observations

The finding in this study of increased norepinephrine levels in patients with CFS seems to be novel, whereas increased levels of epinephrine have been reported sporadically.27,28 A higher level of acute emotional stress among patients with CFS might explain the epinephrine differences; however, plasma levels of norepinephrine are less influenced by such mechanisms.7 Thus, the findings seem to indicate a more substantial alteration of physiology.

A high level of norepinephrine in the antecubital vein

**TABLE 1**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Control Subjects ($N = 57$)</th>
<th>Subjects With CFS ($N = 15$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female gender, n (%)</td>
<td>34 (59.6)</td>
<td>10 (66.7)</td>
</tr>
<tr>
<td>Age, mean (range), y</td>
<td>15.6 (13–18)</td>
<td>15.1 (12–18)</td>
</tr>
<tr>
<td>Weight, mean (range), kg</td>
<td>61.6 (44–99)</td>
<td>59.1 (43–92)</td>
</tr>
<tr>
<td>Height, mean (range), cm</td>
<td>171.4 (149–195)</td>
<td>171.4 (160–192)</td>
</tr>
<tr>
<td>Duration of fatigue, mean (range), mo</td>
<td>NA</td>
<td>31.1 (6–60)</td>
</tr>
</tbody>
</table>

NA indicates not applicable.
plasma might suggest increased sympathetic nerve activity to forearm skin and skeletal muscle. Likewise, a high plasma level of epinephrine might be a result of increased sympathetic nerve activity to the adrenals. However, there are several alternative explanations. Generally, high levels of plasma catecholamines could result from either increased spillover or reduced re- removal, which, in turn, depends on both sympathetic nerve activity, the capacity of different reuptake and breakdown pathways, and local blood flow. Furthermore, a high norepinephrine concentration in forearm venous blood might simply reflect increased arterial levels, which, in turn, could be because of enhanced spillover in other parts of the body.

The plasma levels of metanephrines are not good markers of activity in either the adrenal medulla or the sympathetic neurons and are only weakly correlated with the plasma levels of the respective catecholamines. Thus, similar levels of metanephrines among patients with CFS and control subjects do not rule out a state of catecholamine excess in the former.

The finding of increased TT in patients with CFS is in agreement with previous reports of increased skin temperature in this population but contrasts with 2 other studies that did not find any deviations in core body temperature. However, these latter studies focused primarily on alterations in circadian temperature rhythms. In this study, the increased TT might be partially caused by high levels of epinephrine, which increase basal metabolic rate and heat production. Furthermore, a tendency toward shivering, as reported in our patients with CFS, might contribute.

Blood samples were not obtained from 3 control subjects because of technical difficulties. To reduce the methodologic problem of multiple comparisons, statistical tests were only performed for the variables precooling and Δ cooling.

### TABLE 2
Symptoms Indicative of Thermoregulatory Disturbances

<table>
<thead>
<tr>
<th>Variable</th>
<th>Control Subjects</th>
<th>Subjects With CFS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n/N</td>
<td>%</td>
</tr>
<tr>
<td>Experience of sudden changes in skin color</td>
<td>42/57</td>
<td>73.7</td>
</tr>
<tr>
<td>Redness of skin</td>
<td>37/42</td>
<td>88.1</td>
</tr>
<tr>
<td>Paleness of skin</td>
<td>22/42</td>
<td>52.4</td>
</tr>
<tr>
<td>Experience of shivering hands</td>
<td>8/56</td>
<td>14.0</td>
</tr>
<tr>
<td>Sweating more than others</td>
<td>4/44</td>
<td>9.1</td>
</tr>
</tbody>
</table>

*P ≤ .05, using Fisher’s exact test.
†P ≤ .01, using Fisher’s exact test.
‡This question was added during the course of the study, which explains the lower totals.

### TABLE 3
Catecholamines and Metanephrines Before and After Strong Cooling of 1 Hand

<table>
<thead>
<tr>
<th>Variable</th>
<th>Precooling, Median (95% CI)</th>
<th>Postcooling (10°C Water), Median (95% CI)</th>
<th>Δ Cooling, Median (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control Subjects</td>
<td>Subjects With CFS</td>
<td>Control Subjects</td>
</tr>
<tr>
<td>Norepinephrine, pmol/L</td>
<td>1156 (966 to 1434)</td>
<td>1647* (1478 to 1912)</td>
<td>1494 (1213 to 1807)</td>
</tr>
<tr>
<td>Epinephrine, pmol/L</td>
<td>171 (147 to 192)</td>
<td>216* (161 to 273)</td>
<td>209 (193 to 220)</td>
</tr>
<tr>
<td>Dopamine, pmol/L</td>
<td>130 (112 to 151)</td>
<td>157 (121 to 207)</td>
<td>147 (128 to 172)</td>
</tr>
<tr>
<td>Normetanephrine, nmol/L</td>
<td>0.45 (0.37 to 0.55)</td>
<td>0.36 (0.21 to 0.49)</td>
<td>0.28 (0.23 to 0.40)</td>
</tr>
<tr>
<td>Metanephrine, nmol/L</td>
<td>0.30 (0.27 to 0.34)</td>
<td>0.34 (0.20 to 0.41)</td>
<td>0.25 (0.19 to 0.27)</td>
</tr>
<tr>
<td>3-Methoxytyramine, nmol/L</td>
<td>0.08 (0.07 to 0.09)</td>
<td>0.10 (0.06 to 0.13)</td>
<td>0.08 (0.05 to 0.09)</td>
</tr>
</tbody>
</table>

Blood samples were not obtained from 3 control subjects because of technical difficulties. To reduce the methodologic problem of multiple comparisons, statistical tests were only performed for the variables precooling and Δ cooling.

*P ≤ .01 for differences between groups, using Wilcoxon-Mann-Whitney’s test.
†P ≤ .05 for differences between groups, using Wilcoxon-Mann-Whitney’s test.
### TABLE 4 Thermoregulatory Variables Before and After Moderate Cooling of Left Hand

<table>
<thead>
<tr>
<th>Variable</th>
<th>Precooling (35° C Water), Median (95% CI)</th>
<th>Early Cooling (27° C Water), Median (95% CI)</th>
<th>Δ Cooling 1, (Early Cooling — Precooling), Median (95% CI)</th>
<th>Late Cooling (19° C Water), Median (95% CI)</th>
<th>Δ Cooling 2 (Late Cooling — Precooling), Median (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HR, beats per min</td>
<td>78.7 (67.7 to 73.8)</td>
<td>77.7 (65.2 to 72.9)</td>
<td>-1.6 (-5.1 to 1.0)</td>
<td>73.5 (63.5 to 76.6)</td>
<td>-1.3 (-10.0 to 8.6)</td>
</tr>
<tr>
<td>MBP, mm Hg</td>
<td>81.7 (78.2 to 86.0)</td>
<td>79.2 (78.2 to 86.0)</td>
<td>-2.5 (-5.0 to 1.0)</td>
<td>85.2 (79.2 to 88.8)</td>
<td>2.5 (0.6 to 4.8)</td>
</tr>
<tr>
<td>ASBF left, au</td>
<td>22.2 (117.3 to 371.7)</td>
<td>183.0 (116.8 to 317.1)</td>
<td>-192 (-229 to -152)</td>
<td>245.3 (188.8 to 103.0)</td>
<td>1.4 (-39.0 to 43.7)</td>
</tr>
<tr>
<td>ASBF right, au</td>
<td>22.7 (116.8 to 317.1)</td>
<td>223.1 (116.8 to 317.1)</td>
<td>-192 (-229 to -152)</td>
<td>245.3 (188.8 to 103.0)</td>
<td>1.4 (-39.0 to 43.7)</td>
</tr>
<tr>
<td>TT, °C</td>
<td>35.96 (35.88 to 36.08)</td>
<td>36.21 (36.05 to 36.54)</td>
<td>0.14 (0.02 to 0.22)</td>
<td>36.11 (35.98 to 36.34)</td>
<td>0.08 (-0.02 to 0.22)</td>
</tr>
</tbody>
</table>

To reduce the methodologic problem of multiple comparisons, statistical tests were only performed for the variables precooling and Δ cooling 2. au indicates arbitrary units.

*p ≤ .01 for differences between groups, using Wilcoxon Mann-Whitney test.

*a ≤ .05 for differences between groups, using Wilcoxon Mann-Whitney test.

**Correlation coefficients.

—to increase the postsynaptic sensitivity for norepinephrine.

*Correlation coefficients.

**Correlation coefficients.

—stimulate thermoneuronal activity (Fig 1).
skin areas have found stronger vasodilative responses among patients with CFS than healthy control subjects, suggesting subtle alteration of the endothelial microvascular regulatory system. Other studies have documented a strong relation between CFS and the postural orthostatic tachycardia syndrome, a condition that seems to be characterized by reduced norepinephrine reuptake in the sympathetic synapse. Altogether, the previous studies and our own results underscore the need for research specifically addressing the complicated interaction of adrenergic, cholinergic, and other microvascular control mechanism in patients with CFS.

**Study Limitations**

We instructed the participants to abstain from tobacco products and alcohol/caffeine-containing beverages before the experiments and also asked about alcohol/tobacco/narcotics in the questionnaire without finding any differences between the 2 groups. Still, we cannot completely rule out ingestion of illicit or nonillicit substances that might have influenced the results. Female reproductive hormones exert strong influence on sympathetic cardiovascular regulation; thus, differences among the girls in the 2 groups with regard to sexual development and menstrual cycle might have introduced a bias. However, items addressing menstrual function in the questionnaire did not reveal any significant differences between the 2 groups.

Several symptoms reported by the patients with CFS are not specifically related to thermoregulatory disturbances but could be explained by, for instance, circulatory abnormalities, which indeed have been demonstrated in this group. The neuroendocrine and cardiovascular responses to cooling were addressed in 2 separate experiments. A combined setup might have been more informative; in particular, measuring cardiovascular variables during strong cooling could have yielded better insight into the autonomic responses. Normetanephrine is the product of extraneuronal breakdown of catecholamines, whereas dihydroxyphenylethylene glycol mainly results from intraneuronal metabolism. Concomitant measurement of this metabolite would have given more information on sympathetic nerve function. Between the 2 sets of experiments, the participants were allowed to eat, which might have altered skin blood flow. However, because the meals were standardized, this could hardly account for the observed differences between the 2 groups. In the moderate-cooling experiment, ASBF and TT were not measured before the acclimatization period; thus, the findings are not necessarily valid for ordinary ambient temperatures.

This study used a healthy control group, having a higher level of activity than the patient group. Thus, it is possible that our findings are a mere consequence of inactivity rather than of the underlying CFS pathophysiology. Additional research projects should consider using sedentary healthy control subjects or patients with other diseases having a comparable activity level to the patients with CFS. Finally, only 15 adolescent patients with CFS were studied, reducing the statistical power and bringing into question the generalizability of the results.

**CONCLUSIONS**

Taken together, our results suggest that adolescent patients with CFS have abnormal catecholaminergic-dependent thermoregulatory responses both at rest and during local skin cooling. These results seem to support a hypothesis of sympathetic dysfunction in CFS. Furthermore, they might explain important clinical symptoms.

**ACKNOWLEDGMENTS**

Dr Wyller has received a research grant from the University of Oslo.

We thank Elisabeth Getz (Department of Pediatrics, Rikshospitalet-Radiumhospitalet Medical Center) and Torun Flatebo (Department of Physiology, University of Oslo) for technical assistance during the experiments and Helene Gjone (Department of Child Psychiatry, Rikshospitalet-Radiumhospitalet Medical Center) for assessment of patients with CFS.

**REFERENCES**

8. Royal College of Paediatrics and Child Health. Evidence Based Guidelines for the Management of CFS/ME (Chronic Fatigue Syndrome/Myalgic Encephalopathy) in Children and Young Adults. London, United Kingdom: Royal College of Paediatrics and Child Health; 2004


Use of Complementary and Alternative Medicine in a General Pediatric Clinic

Dany Jean, MD, Claude Cyr, MD, FRCPC

Département de Pédiatrie, Faculté de Médecine et des Sciences de la Santé, Université de Sherbrooke, Sherbrooke, Québec, Canada

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

ABSTRACT

BACKGROUND. Use of complementary and alternative medical therapies is common and increasing, particularly for children with chronic disease.

OBJECTIVES. The purpose of this work was to describe the use of complementary and alternative medicine by children and to identify factors that may influence the use of complementary and alternative medicine.

PATIENTS AND METHODS. We conducted a cross-sectional descriptive study with children who were visiting a pediatric outpatient clinic. Parent’s satisfaction about primary care was evaluated with the Parent’s Perceptions of Pediatric Primary Care Quality questionnaire.

RESULTS. Fifty-four percent of children used ≥1 type of complementary and alternative medicine in the previous year. No sociodemographic characteristic difference was found between user and nonuser groups. Children most often used complementary and alternative medicine to treat musculoskeletal problems (27%), psychological problems (24%), or infections (20%). Factors that influenced complementary and alternative medicine use were “word of mouth” (36%), “reference by a physician” (28%), “personal experience by parents” (28%), and “no adequate resources in ‘traditional’ medicine” (21%). Forty-seven percent of complementary and alternative medicine users used prescribed medications simultaneously. Most users (75%) believed that complementary and alternative medicine had no potential adverse effects or interactions with prescribed medication. Only 44% of complementary and alternative users were known as such by their physician. The primary care satisfaction was significantly lower in complementary and alternative users versus nonusers. Parents of complementary and alternative users were less satisfied in the areas of accessibility, knowledge of the patient, and communication.

CONCLUSIONS. Complementary and alternative medicine was used by 54% of the children in our cohort. Complementary and alternative medicine users were less satisfied with primary care than nonusers. Only 44% of complementary and alternative medicine users were known by their physician. It is important that
physicians systematically elicit families’ expectations of treatment and be aware of the range of therapies used by children.

The use of alternative medicine in North America is increasing.1-4 Many studies evaluating the use of complementary and alternative medicine (CAM) by children with chronic illnesses, such as cystic fibrosis, juvenile arthritis, cancer, and asthma, found a user rate as high as 72%.5-7 A classic study by Spiegelblatt et al3 showed that 11% of Canadian children seen in a pediatric outpatient clinic consulted ≥1 CAM practitioner. Similar studies on the use of CAM in children who were not chronically ill showed prevalence ranging between 2% and 55%.2,8,9

The mainstream medical population seems to manifest interest for CAM. Physicians want to know more about CAM and occasionally encourage their patients to use them.10 Sikand et al11 showed that 84% of pediatricians believed that some of their patients used CAM, and most of the physicians thought that this population represented <10% of their patients. Approximately 50% of families had discussed with their physician their use of CAM. However, 76% of physicians believed their patients would tell them if they were using CAM.11 Studies have also demonstrated that most adult patients do not inform their physicians that they are using CAM.

Adverse effects are possible with CAM, and multiple pharmacologic interactions exist with prescription medication.4 These facts demonstrate the importance for physicians to be knowledgeable about CAM and to be aware of their patient use of CAM.

The aims of this study were to evaluate simultaneously the use of CAM by children, to evaluate the awareness of physicians about this use, and to compare satisfaction with primary care between CAM users and nonusers.

METHODS
This cross-sectional descriptive study was conducted using a convenience sample of parents of patients presenting in a pediatric outpatient clinic at a university-affiliated general hospital in Estrie (Quebec, Canada) over a 4-week period (October and November 2003). None of the health care providers working in this clinic were practitioners of any form of CAM. CAM included: chiropractic remedies, naturopathy (dietary supplements, medicinal plants, and dietary manipulation), homeopathy, massage, acupuncture, Reiki/energy care (therapy involving the use of energy fields including biofield therapies and bioelectromagnetic-based therapies), hypnosis, osteopathy (any form of osteopathic manipulation), folk remedies, and bone setting (any hands-on techniques to alleviate pain, restore function, or promote health and well being). Any other forms of CAM not listed above were sought and included if disclosed. We used 2 questionnaires: one for patients/families and the other for physicians. Parents and physicians were told that the study was being conducted to examine the types and quality of health care received by the child. Care was taken so that only 1 questionnaire was completed for each child, and all of the collected data remained confidential and anonym.

The patient/family questionnaire was composed of 39 questions: general information about the child, information about CAM use, and a section with the French translation of the Parent’s Perceptions of Pediatric Primary Care Quality (P3C) questionnaire. This questionnaire measures the parents’ degree of satisfaction on ~6 quality domains of primary care: longitudinal continuity, access, contextual knowledge, communication, comprehensiveness, and coordination. Computing the mean of the nonmissing values on each scale formed the total score, as well as scores for each subscale. Scores ranged from 0 to 100, with 100 being best. The P3C questionnaire has good face validity, good reliability (internal consistency with Cronbach’s coefficient of .95 for total score and between .75 and .95 for subscales) and construct validity (convergent and divergent validity).12 Parents were asked whether the child’s doctor had been informed about their use of CAM. The physician questionnaire included questions about the perceived satisfaction with health care and the types of therapies used by the patient, including CAM.

Statistical analysis was performed using StatView 5.0 (SAS Institute Inc, Cary, NC) to provide descriptive statistics. Nominal data were analyzed with χ² test and ordinal data with the Mann-Whitney U test. The study was approved and conducted in accordance with the ethical standards set by the research ethics committee.

RESULTS
On 200 questionnaires distributed, 114 pairs were completed (66%). The patient/family questionnaires were principally completed by mothers (84%). Almost half of the parents had completed a college degree or more. The majority of parents lived as a couple (83%), were employed (86%), and had familial income more than $20 000 per year (79%). The primary care provider for children was a pediatrician in 56% of instances. Fifty-four percent (61 of 114) of the respondents had used ≥1 CAM for their children in the year before the study. We found no sociodemographic difference between the groups of CAM users and nonusers (see Table 1).

The most frequently used CAM types are listed in Table 2. The most frequent health problems that justified CAM use were musculoskeletal problems (27%), psychological problems (24%), infections (20%), asthma/allergies (15%), pain (8%), skin problems (8%), and colic (8%). Fifty-two percent of children were using a prescription medication at the time of the survey. 36%
The patients used CAM for a variety of health issues, principally for musculoskeletal, psychological, and infectious problems. Conventional medicine offers often-limited solutions for these chronic problems. Word-of-mouth was the most common factor influencing a caretaker’s decision to use such a therapy. Many caretakers reported >1 factor influencing their decision. Families may choose a CAM therapy because it will not require a visit to a doctor or a clinic and because they perceive the therapy as more natural. A common explanation for the increasing popularity of CAM is the dissatisfaction with primary care.

Our study population might represent more chronically ill children than the general population, because the patients were followed in a pediatric clinic (secondary or tertiary care). This might explain a higher prevalence than studies in primary care but lower than studies with children with chronic illness.

The method used (questionnaires) can create a selection bias, because the CAM users group could be overestimated. The use of CAM and questions about reasons and motivations are subject to recall bias, because it is a retrospective account over the past year. Potential limitations to our study also include small sample size. Our results may not be representative of the true prevalence of CAM use in the pediatric population. We also used only a single Canadian institution for our survey, which would fail to account for any regional or geographic differences in CAM usage among the pediatric population.

Although we found a high prevalence of CAM usage in our patient population, 53% of CAM users reported that they did not inform their health care provider that they were using CAM. We also found moderate rates of combining prescription medicines and CAM, risking adverse events. Eighteen percent of our population simultaneously used herbal medicines and prescription medication, and 75% believed that CAM had no potential adverse effect or interaction. This raises important issues of safety, especially in view of the poor communication shown by parents and physicians about CAM. These findings are not new and support previous studies of poor CAM disclosure rates by parents with pediatricians. Some CAM treatments may be associated with adverse effects or interactions with conventional therapies. Given the increasing use of CAM and the significant degree of underreporting demonstrated in our study, as well as in others, it is important for health care providers to actively question parents and patients on possible CAM use.

In our study, CAM users seemed less satisfied with primary care quality than nonusers as measured with the P3C questionnaire. One interpretation might be that dissatisfaction with “conventional” medicine lead to the use of CAM. On the other side, interest for these therapies by patients (with their individual values and be-

### Table 1: Characteristics of CAM Users Versus Nonusers

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>CAM Users</th>
<th>Nonusers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age father (mean), y</td>
<td>39</td>
<td>38</td>
</tr>
<tr>
<td>Age mother (mean), y</td>
<td>36</td>
<td>35</td>
</tr>
<tr>
<td>Age child (mean), y</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>No. of children in family (median)</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Mother works as a health care professional, %</td>
<td>28</td>
<td>15</td>
</tr>
<tr>
<td>Father works as a health care professional, %</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Marital status, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Couple (never separated)</td>
<td>75</td>
<td>69</td>
</tr>
<tr>
<td>Reconstituted family</td>
<td>12</td>
<td>19</td>
</tr>
<tr>
<td>Single-parent</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>Vaccination up to date, %</td>
<td>92</td>
<td>96</td>
</tr>
<tr>
<td>Child currently taking prescription medication, %</td>
<td>48</td>
<td>43</td>
</tr>
<tr>
<td>Child with chronic illness, %</td>
<td>25</td>
<td>12</td>
</tr>
</tbody>
</table>

### Table 2: Prevalence of CAM Use

<table>
<thead>
<tr>
<th>CAM Type</th>
<th>Prevalence, %</th>
<th>Past 12 mo</th>
<th>Lifetime Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homeopathy</td>
<td>30</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>Naturopathy</td>
<td>20</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Chiropractic remedy</td>
<td>19</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Osteopathy</td>
<td>13</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Massage therapy</td>
<td>12</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Folk remedies</td>
<td>9</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Reiki</td>
<td>8</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>“Bone setter”</td>
<td>6</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Acupuncture</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Hypnosis</td>
<td>0</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

used natural products like herbal remedies, and 19% used both simultaneously. Seventy-five percent of the patients/families did not believe that CAM could have adverse effects or interactions with medication.

CAM users were less satisfied with the primary care received by their children than nonusers according to the P3C (global score, users versus nonusers, mean ± SD: 74 ± 16 vs 80 ± 15; P = .01). Satisfaction was particularly lower in the CAM users group in the areas of accessibility, contextual knowledge, and communication. Forty-seven percent of CAM users had discussed CAM with their physicians. Only 44% of actual CAM users were identified as such by their physician.

**DISCUSSION**

The prevalence of CAM use in this study (54%) is higher than reported previously in a similar setting but in the range of recent reports about chronically ill children. Our survey used a broader definition of CAM and included practices such as folk remedies. These remedies may be easier to access than those identified by Spigelblatt et al., and that might explain the higher rate of CAM usage in our study. The most popular CAM types were homeopathy, chiropractic remedies, and naturopathy. These results were similar to many others studies.

The method used (questionnaires) can create a selection bias, because the CAM users group could be overestimated. The use of CAM and questions about reasons and motivations are subject to recall bias, because it is a retrospective account over the past year. Potential limitations to our study also include small sample size. Our results may not be representative of the true prevalence of CAM use in the pediatric population. We also used only a single Canadian institution for our survey, which would fail to account for any regional or geographic differences in CAM usage among the pediatric population.
rients) lead them to be less satisfied by the conventional medical context. One of the frequent reasons for CAM use was persistent medical problems that were perceived not to have improved with conventional medical treatment. This reason, combined with a general dissatisfaction with conventional medicine, accounted for one third of the total of reasons given. Two US studies concluded that the use of CAM therapies cannot be attributed primarily to perceived dissatisfaction with conventional medical care or caregivers and that many adults seek, explore, and experience benefits from both conventional and CAM therapies. This contrasts with previous comments suggesting that the high prevalence of CAM use largely represents a societal rejection of orthodox and conventional medical care. If health care professionals are to effectively support individuals in making informed, safe, and appropriate choices, it is critical that they develop greater awareness of the nature of, potential efficacy of, and reasons for CAM use. Given the frequency of use of a variety of CAM therapies, caregivers should inquire about CAM usage in all children at each and every office visit. Further studies about CAM use should look into the reasons for use and incorporate the perception of parents about the quality of primary care.

CONCLUSIONS

The use of CAM by children is frequent. Poor communication about CAM use leads to lack of knowledge by physicians. CAM users are less satisfied with primary care quality, especially for aspects of communication, contextual knowledge, and accessibility. Physicians should have a working knowledge of the escalating literature on CAM to be in a position to discuss implications of use. Because they are being used to treat children, physicians who care for children should be aware of the various types of CAM therapies, scientific evidence supporting or refuting their use, and the potential adverse effects of each therapy.

REFERENCES

8. Crawford NW, Cincotta DR, Lim A, Powell CV. A cross-sectional survey of complementary and alternative medicine use by children and adolescents attending the University Hospital of Wales. BMC Complement Altern Med. 2006;6:16
Frequency and Characteristics of Pediatric and Adolescent Visits in Naturopathic Medical Practice

Wendy Weber, ND, MPH\textsuperscript{a}, James A. Taylor, MD\textsuperscript{b}, Rachelle L. McCarty, ND, MPH\textsuperscript{a}, Amy Johnson-Grass, ND, MSN\textsuperscript{c}

\textsuperscript{a}School of Naturopathic Medicine, Bastyr University, Kenmore, Washington; \textsuperscript{b}Child Health Institute, University of Washington and Children’s Hospital and Regional Medical Center, Seattle, Washington; \textsuperscript{c}Minnesota Consortium for CAM Clinical Research, University of Minnesota, Minneapolis, Minnesota

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

ABSTRACT

OBJECTIVES. This work sought to identify naturopathic physicians in Washington State who frequently provide pediatric care and to describe the conditions treated and therapies recommended for children.

PATIENTS AND METHODS. A mailed survey of licensed naturopathic physicians residing in Washington State collected demographic information and practice descriptions. For naturopathic physicians treating \( \geq 5 \) pediatric patients per week, data were collected on the conditions seen and treatments provided to children during a 2-week period.

RESULTS. Of 499 surveys delivered to providers, 251 surveys were returned (response rate: 50.3%). Among the 204 naturopathic physicians currently practicing, only 31 (15\%) saw \( \geq 5 \) children per week. For these pediatric naturopathic physicians, pediatric visits constituted 28\% of their office practice. Pediatric naturopathic physicians were more likely to be licensed midwives (19.4\% vs 0.6\%) and treated significantly more patients per week (41.6 vs 20.2) than naturopathic physicians who provided less pediatric care. Eighteen of the 31 pediatric naturopathic physicians returned data on 354 pediatric visits; 30.5\% of the visits were by children <2 years old, and 58.5\% were by those <6 years old. The most common purpose for presentation included health supervision visits (27.4\%), infectious disease (20.6\%), and mental health conditions (12.7\%). Pediatric naturopathic physicians provided immunizations during 18.6\% of health supervision visits by children <2 years old and 27.3\% of visits by children between the ages of 2 and 5 years.

CONCLUSIONS. Although most naturopathic physicians in Washington treat few children, a group of naturopathic physicians provide pediatric care as a substantial part of their practice. Based on the ages of children seen and the conditions treated, pediatric naturopathic physicians may provide the majority of care for some children. Efforts should be made to enhance collaboration between naturopathic physicians and conventional providers so that optimal care can be provided to children.
The practice of naturopathic medicine is licensed in 14 states in the United States, as well as the District of Columbia, Puerto Rico, and the US Virgin Islands. Under these licenses, naturopathic physicians provide care to infants, children, adolescents, and adults. Naturopathic physicians (NDs) are trained to treat the whole person with natural means, such as nutrition, exercise, lifestyle modification, vitamins, herbal medicines, nutritional supplements, and hydrotherapy. In addition to these modalities, in some states, licensed NDs can prescribe conventional medications, including antibiotics and vaccines. Licensing requires graduation from 1 of 5 accredited naturopathic medical schools in the United States or Canada (1 additional school is a candidate for accreditation). NDs receive a minimum of 4100 hours of education, including 1200 hours of clinical training.

The majority of patients seen by NDs are adults, and most visits by adults are for the treatment of chronic conditions, such as fatigue, back symptoms, and headache. There are limited data on naturopathic medicine use by children. A survey of 15 NDs in Massachusetts reported that 19% of visits to an ND were by children and that, on average, 5 pediatric patients were seen per week. In another study, 99 NDs provided visit data on 20 consecutive patients; 10% to 12% of these visits were made by pediatric patients <15 years of age. Wilson et al summarized the chief complaints reported by 482 parents bringing their child in for naturopathic care at a large academic naturopathic clinic in Toronto, Ontario, Canada. The authors reported that 23% of pediatric patients presented with skin disorders, 17% with gastrointestinal complaints, and 15% with psychiatric or behavioral disorders. In a review of insurance claims data, Bellas et al found that 1% of children enrolled in 2 large insurance plans in Washington State had visited an ND. These reports provided few data on the care provided by NDs to children.

To better characterize pediatric naturopathic care, we conducted a survey of NDs in Washington State. A main goal of the study was to identify NDs who provided care to a substantial number of children and to compare these practitioners to NDs who saw fewer pediatric patients. A second goal of the project was to describe the chief complaints from pediatric visits to NDs and the categories of treatments provided.

METHODS
In June 2004, the Department of Health in Washington State was contacted and provided the investigators with a list of NDs licensed in Washington. A cover letter and survey were mailed to all of the NDs with mailing addresses in the state of Washington in July 2004. A $1 incentive was included with each survey. For any survey returned as undeliverable, the online databases of the American Association of Naturopathic Physicians and the Washington Association of Naturopathic Physicians were searched in attempt to find an updated address for the ND (www.naturopathic.org and www.wanp.org). A second mailing was sent in November 2004 to all of the NDs who did not return the initial survey, along with another $1 incentive.

Contacted NDs were asked to complete a 1-page survey that included demographic items and questions about their practices. Demographic information included the age, gender, race, and ethnicity of the provider; information about other types of practitioners working in the office; an estimate of the average number of children or adolescents treated per week; and whether the provider’s practice marketed for pediatric patients. NDs who reported an average of ≥5 pediatric patient visits per week were asked to record information including the age, gender, chief complaints, and treatment provided for each child <18 years old who was seen in the office during a 2-week time period. Surveys and visit data were collected from July 2004 to January 2005.

For the analysis, characteristics of NDs who estimated that they saw ≥5 pediatric patients per week were compared with those of respondents who saw fewer children. Differences in the number of years in practice, number of patients treated per week, and proportion of pediatric patients seen were assessed with t tests. χ² tests were used to evaluate the significance of the differences between NDs who reported seeing substantial numbers of children in their practices and those who saw fewer children for types of additional degrees, types of providers sharing practices, and specific marketing for pediatric health care services. Differences were considered significant if the P value was <.05.

Data on patient visits were aggregated by age of the patient (<2, 2–5, 6–11, and 12–17 years old) and by primary diagnosis. Major diagnostic categories included health supervision visits (for visits characterized as “well-child examination” or “screening physical”), upper respiratory tract infections (cold, cough, and otitis media), allergies, eczema, attention-deficit/hyperactivity disorder (ADHD), and autism. Categories of treatments were further evaluated by specific diagnoses and within selected age groups.

This study was approved by the Bastyr University Institutional Review Board. Language was included in the cover letter stating that participation was voluntary, and consent was implied if the completed survey was returned.

RESULTS
The names and addresses of a total of 649 NDs were provided by the Washington State Department of Health. An out-of-state mailing address was listed for 136 of these NDs, excluding them from the study. Surveys were mailed to the remaining 513 licensed NDs residing in the state of Washington. Fourteen surveys (2.7%) were returned to sender as nondeliverable, and...
no correct forwarding address was available. Of the 499 surveys delivered to NDs, 251 completed surveys were returned (50.3%).

Returned surveys indicated that 18.7% of NDs were not actively in practice, and demographic data were not collected on these individuals. Of the 204 practicing NDs who completed the survey, 31 (15.2%) reported treating ≥5 pediatric patients per week (termed “pedsNDs”). In Table 1, the characteristics of these providers and their practices are compared with those of NDs who reported fewer visits with children (termed “nonpedsNDs”). Seven completed surveys did not indicate the number of pediatric patients treated per week; therefore, data from these surveys were not included in the comparative analysis between pedsNDs and nonpedsNDs. There were few significant differences in provider characteristics between pedsNDs and nonpedsNDs. However, although the overall proportion with additional degrees was similar, pedsNDs were more likely to also be licensed midwives (19.4% vs 6.0%; P < .001). In terms of their practices, pedsNDs reported significantly more visits per week and were more likely to specifically market pediatric health care services. Overall, 78.9% of all of the NDs reported sharing their office with other providers including the following: other NDs (65.2%), licensed acupuncturists (40.4%), massage therapists (40.4%), chiropractic doctors (18.0%), MDs (13.7%), PhDs (6.8%), counselors (6.2%), registered nurses (5.0%), licensed midwives (4.4%), and nutritionists (4.4%). PedsNDs were less likely to have shared their office with an acupuncturist (17.9% vs 44.6%; P < .01) and more likely to share their office with a licensed midwife (17.9% vs 1.4%; P < .001) than those who saw fewer children.

Patient visit data were received from 18 (58%) of the 31 pedsNDs; information on 354 visits was collected. Of the patients seen by pedsNDs, 30.5% of the pediatric patients were under the age of 2 years; 28.0% of the patients were between 2 and 5 years of age; 22.6% were between 6 and 11 years of age; and 18.9% were between 12 and 18 years of age. The mean age of children and adolescents visiting pedsNDs was 5.8 years (SD: 5.4); the median age was 4 years. Nearly 45% of visits were by girls, and 55% of visits to pedsNDs were by boys.

The most common chief complaints recorded by pedsNDs for the patients’ visits included health supervision visits (27.4%), upper respiratory infections (URIs) (18.4%; including URI [7.9%], otitis media [5.9%], and cough [4.5%]), allergies (6.2%), eczema (4.2%), ADHD (3.7%), and autism (2.5%). An estimated 24.0% (95% confidence interval: 19.7%–28.8%) of visits were for chronic conditions, including allergies, asthma, eczema, and mental health conditions. Table 2 provides details of the treatments recommended by the pedsNDs for each of the conditions listed above. Except for children with autism, pedsNDs recommended >1 category of treatment at 46% of visits.

Within each age group, the 5 most common reasons for visiting a pedsND included health supervision visits, upper respiratory tract infection, allergies, skin disorders, and mental health conditions (Table 3). In addition to the reasons noted above, 9.3% of visits by children <2 years old were for infant conditions (colic, teething, failure to thrive, or thrush). Immunizations were given during 18.6% of health supervision visits by children <2 years of age. Respiratory conditions (asthma or bronchitis) were recorded as the chief complaint for 61.1% of children ages 2 to 5 years. In this age group, immunizations were administered in 27.3% of health supervision visits. Oral antibacterial or antifungal prescriptions were prescribed in 6.1% of visits; however, no antibiotics were prescribed for children presenting with upper respiratory symptoms. Other reasons for visits in adolescents ages 12 to 18 years included 9% for gastrointestinal complaints (diarrhea or irritable bowel syndrome) and 7.5% for gynecologic disorders (amenorrhea, vaginitis, or premenstrual syndrome). Overall, no antibiotics were prescribed for any of the 65 children presenting with upper respiratory symptoms.

### DISCUSSION
The results of this survey indicate that most NDs in Washington State see limited numbers of children in their practices. However, there is a distinct group of NDs, representing 15% of survey respondents, for whom pediatric care is a substantial part of their practice. These

---

**TABLE 1**

Demographics of NDs in Active Clinical Practice Reporting That They Treat ≥5 Pediatric Patients per Week and Those Who Do Not

<table>
<thead>
<tr>
<th>Variable</th>
<th>PedsNDs (N = 31)</th>
<th>NonpedsNDs (N = 166)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female NDs, %</td>
<td>63.3</td>
<td>65.4</td>
</tr>
<tr>
<td>Age category of NDs, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21–30 y</td>
<td>9.7</td>
<td>7.2</td>
</tr>
<tr>
<td>31–40 y</td>
<td>35.5</td>
<td>34.3</td>
</tr>
<tr>
<td>41–50 y</td>
<td>35.5</td>
<td>33.1</td>
</tr>
<tr>
<td>51–60 y</td>
<td>19.4</td>
<td>19.9</td>
</tr>
<tr>
<td>61–70 y</td>
<td>0</td>
<td>1.8</td>
</tr>
<tr>
<td>&gt;70 y</td>
<td>0</td>
<td>0.6</td>
</tr>
<tr>
<td>Not reported</td>
<td>0</td>
<td>3.0</td>
</tr>
<tr>
<td>White, %</td>
<td>87.1</td>
<td>78.9</td>
</tr>
<tr>
<td>Additional degrees, %</td>
<td>35.5</td>
<td>36.1</td>
</tr>
<tr>
<td>Mean No. of years in practice (SD)</td>
<td>11.2 (7.7)</td>
<td>9.0 (7.5)</td>
</tr>
<tr>
<td>Group practice, %</td>
<td>90.3</td>
<td>79.3</td>
</tr>
<tr>
<td>Mean No. of partners in practice (SD)</td>
<td>2.8 (2.0)</td>
<td>5.7 (9.3)</td>
</tr>
<tr>
<td>Mean No. of visits per week (SD)</td>
<td>41.6 (34.5)</td>
<td>202 (185)</td>
</tr>
<tr>
<td>Mean No. of pediatric visits per week (SD)c</td>
<td>9.5 (8.7)</td>
<td>1.5 (1.1)</td>
</tr>
<tr>
<td>Mean percentage of pediatric patients (SD)</td>
<td>28.0 (18.5)</td>
<td>11.6 (14.5)</td>
</tr>
<tr>
<td>Market for pediatric patients, %</td>
<td>22.6</td>
<td>8.1c</td>
</tr>
</tbody>
</table>

* P < .001.

b Statistical analysis was not performed, because this defines the difference between the 2 groups.

c P < .05.
were seen by NDs (45% and 55%, respectively). Sim-
we found that nearly equal numbers of girls and boys
find that women are more likely to use CAM treatments,
and alternative medicine (CAM), which consistently
midwifery care themselves.
and often partner with licensed midwives or provide
pedsNDs see almost 10 children per week in their offices,
may specifically market the pediatric care they provide,
and often partner with licensed midwives or provide
midwifery care themselves.
Contrary to reports of adult use of complementary
and alternative medicine (CAM), which consistently
find that women are more likely to use CAM treatments,
we found that nearly equal numbers of girls and boys
were seen by NDs (45% and 55%, respectively). Similar
adults who are seen by NDs, 24% of the visits by
pediatric patients to pedsNDs were for the treatment of
chronic disorders, particularly those in which therapies
provided by conventional practitioners are either
controversial to some parents, such as with ADHD, or of
limited benefit, as with autism. However, in our study,
the 2 most common reasons for visits by children to
pedsNDs were health supervision and URI symptoms.
Preschool-aged children made most of these visits.
Health supervision and URI symptoms are also the most
common reasons that children see conventional provid-
ers. These data suggest that pedsNDs provide the
majority of care for a group of pediatric patients.
Our finding that a group of children receive much of
their care from NDs has implications for the health care
provided to pediatric patients, as well as providing an
opportunity for collaboration between conventional and
alternative practitioners. The most striking example of
this is immunizations. Despite reports that NDs are op-
posed to vaccination, we found that immunizations
were administered at 18.6% of health supervision visits
for children <2 years old and at 27.3% of health
supervision visits for children 2 to 5 years old. Based on the
recommended periodicity for health supervision visits to
pediatricians and the current immunization schedule, it
is likely that vaccines are provided at >50% of health
supervision visits for children <2 years old to conven-
tional providers. This suggests that there may be am-
ivalence regarding vaccinations among pedsNDs and/or
the parents of their patients. In a previous survey, 1 of 15
NDs in Massachusetts opposed immunization, 3 of 15
actively recommended immunization, and the remain-
der did not actively make recommendations. It is likely
that parents who oppose vaccination seek out pedsNDs
as primary care providers for their children. Salmon et
al found that parents of children who were exempted
from school immunization requirements had signifi-
cantly more negative beliefs about vaccines, and 11.5%
received primary health care from an alternative pro-
vider versus 0.3% of those who were fully immunized.
Currently, NDs are licensed to administer vaccines in a
limited number of states, so children receiving care ex-
clusively from NDs in unlicensed states may have limited
access to immunization services.
Although NDs are licensed to prescribe antibiotics in
Washington state, 0 of 65 children seen for URI symp-
toms by a pedsND were prescribed these medications.
Because the upper limit of the 95% confidence interval
for a point estimate of 0% with a sample of 65 children is
4.6%, our results suggest that <5% of pediatric
patients presenting to the pedsNDs in this study with
URI are prescribed antibiotics. This is in contrast to the
experience of young children seen by pediatricians be-
cause of URI symptoms. In a study of community pedi-
atricians in the Seattle, WA, area, ≥1 antibiotic was
prescribed during 46% of visits by patients <3 years old
who presented with cough and cold symptoms. It is
possible that, by adopting some of the techniques used

### Table 2: Percentage of Visits With Treatment Recommended and Mean Age for Specific Conditions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Health Supervision</th>
<th>URI</th>
<th>Allergies</th>
<th>ECZema</th>
<th>ADHD</th>
<th>Autism</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 97)</td>
<td>(n = 65)</td>
<td>(n = 22)</td>
<td>(n = 15)</td>
<td>(n = 13)</td>
<td>(n = 9)</td>
</tr>
<tr>
<td>Herbal, %</td>
<td>9.5</td>
<td>52.3</td>
<td>18.2</td>
<td>26.7</td>
<td>7.7</td>
<td>0</td>
</tr>
<tr>
<td>Nutrition, %</td>
<td>59.8</td>
<td>27.7</td>
<td>36.4</td>
<td>66.7</td>
<td>38.5</td>
<td>22.2</td>
</tr>
<tr>
<td>Vitamin, %</td>
<td>24.7</td>
<td>26.2</td>
<td>13.6</td>
<td>26.7</td>
<td>23.1</td>
<td>22.2</td>
</tr>
<tr>
<td>Homeopathy, %</td>
<td>3.1</td>
<td>52.3</td>
<td>77.3</td>
<td>33.3</td>
<td>69.2</td>
<td>88.9</td>
</tr>
<tr>
<td>Physical medicine, %</td>
<td>0</td>
<td>7.7</td>
<td>0</td>
<td>6.7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Supplements, %</td>
<td>1.0</td>
<td>3.1</td>
<td>22.7</td>
<td>20.0</td>
<td>0</td>
<td>11.1</td>
</tr>
<tr>
<td>Lifestyle, %</td>
<td>32.3</td>
<td>10.8</td>
<td>13.6</td>
<td>20.0</td>
<td>38.5</td>
<td>0</td>
</tr>
<tr>
<td>Multiple treatments, %</td>
<td>46.4</td>
<td>52.3</td>
<td>59.1</td>
<td>66.7</td>
<td>46.2</td>
<td>22.2</td>
</tr>
<tr>
<td>Other treatments, %</td>
<td>22.7</td>
<td>15.4</td>
<td>22.7</td>
<td>13.3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>No treatment, %</td>
<td>17.5</td>
<td>3.1</td>
<td>0</td>
<td>6.7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Vaccine administered, %</td>
<td>17.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Antifungal/antibiotic medication, %</td>
<td>2.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Age of patients, mean (SD), y</td>
<td>2.9 (4.0)</td>
<td>3.9 (3.2)</td>
<td>10.7 (4.3)</td>
<td>4.6 (4.3)</td>
<td>9.2 (3.9)</td>
<td>7.2 (5.5)</td>
</tr>
</tbody>
</table>

### Table 3: Percentage of Visits for Specific Condition Within Different Age Groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>&lt;2 y (n = 108)</th>
<th>2–5 y (n = 99)</th>
<th>6–11 y (n = 80)</th>
<th>12–18 y (n = 67)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health supervision, %</td>
<td>54.6</td>
<td>22.2</td>
<td>11.3</td>
<td>9.0</td>
</tr>
<tr>
<td>URI, %a</td>
<td>15.7</td>
<td>30.3</td>
<td>18.8</td>
<td>4.5</td>
</tr>
<tr>
<td>Allergies, %</td>
<td>0</td>
<td>3.0</td>
<td>10.0</td>
<td>16.4</td>
</tr>
<tr>
<td>Skin conditions, %b</td>
<td>11.1</td>
<td>10.1</td>
<td>11.3</td>
<td>7.5</td>
</tr>
<tr>
<td>Mental health, %c</td>
<td>2.8</td>
<td>11.1</td>
<td>20.0</td>
<td>14.9</td>
</tr>
</tbody>
</table>

a Data include URI, cold, cough, or otitis media.
b Data include eczema, rash, or skin problems.
c Data include ADHD, autism, depression, or anxiety.
by the pedsNDs participating in our study, conventional practitioners might reduce the injudicious use of antibiotics in children with cold and cough symptoms.

The generalizability of the survey findings to NDs across the United States and internationally may be limited, because this survey was conducted in a licensed state. It is possible that the conditions and treatments reported by NDs in the survey may not be the same as those practicing in states where NDs are not licensed. Furthermore, Washington state law has required many insurance companies to reimburse for services provided by NDs.16 Cherkin et al1 reported that 50% of visits to Washington NDs were covered by insurance in 1998–1999. The results of a recent study analyzing Washington state insurance claims indicate that NDs see a similar spectrum of conditions as conventional practitioners.6 Lack of insurance coverage may change the reasons that parents bring their children in for treatment by NDs. For example, in states with a lower frequency of insurance reimbursement for ND care, parents may be less likely to pay out-of-pocket for a health supervision visit by an ND when it would be covered if performed by an MD. Thus, the results of this survey may only generalize to NDs practicing in licensed states with high levels of medical insurance reimbursement for this care.

Even in Washington, a state in which NDs are licensed, insurance coverage is available, and a naturopathic medical school is located, the results of this study indicate that pediatric care by most NDs is limited. However, in addition to adjunct care for chronic or complicated conditions, our results strongly suggest that there is a group of children for whom pedsNDs may be the sole health care providers. Some of these pedsNDs are providing immunizations to their patients, yet not at the same frequency as pediatricians. Research is needed to further elucidate the beliefs and recommendations of NDs regarding vaccines and the beliefs and preferences of the parents who bring their child to these providers. In addition, studies examining the practice of NDs in unlicensed states with limited or no insurance coverage are needed to determine whether these providers treat as many pediatric patients with similar conditions. Nationally, interest and use of CAM is increasing; it is likely that the demand for NDs as pediatric primary care providers will also increase.

ACKNOWLEDGMENTS
This publication was made possible by National Center for Complementary and Alternative Medicine grants K23AT000929 and T32 AT000815 and a faculty seed grant from Bastyr University.

REFERENCES
Cardiovascular Effects of Sibutramine in the Treatment of Obese Adolescents: Results of a Randomized, Double-Blind, Placebo-Controlled Study

Stephen R. Daniels, MD, PhD, Barbara Long, MD, MPH, Scott Crow, MD, Dennis Styne, MD, Melinda Sothern, PhD, Ileana Vargas-Rodriguez, MD, Lisa Harris, MD, Julia Walch, BS, Olga Jasinsky, BS, MBA, Kristine Cwik, RN, MBA, Ann Hewkin, MSc, Vicky Blakesley, MD, PhD, for the Sibutramine Adolescent Study Group

ABSTRACT

BACKGROUND. Adolescent obesity is a major public health problem. Treatment options in addition to behavioral therapy could include pharmacotherapy with sibutramine.

OBJECTIVES. Concerns regarding increases in blood pressure and heart rate after sibutramine treatment in some adult patients precipitated the present analysis, which evaluated the cardiovascular safety of sibutramine plus a behavioral therapy program in obese adolescents.

PATIENTS AND METHODS. With this 12-month, randomized, double-blind, placebo-controlled trial in 33 US clinics we studied 498 adolescents aged 12 to 16 years with multiethnic backgrounds and BMIs of 28.1 to 46.3 kg/m².

RESULTS. The subjects were randomly assigned to behavioral therapy plus 10 mg of sibutramine or behavioral therapy plus placebo daily. At the end point, there was a mean treatment group difference in BMI of 2.6 kg/m² in favor of sibutramine. Small mean decreases in blood pressure and pulse rate were seen in both sibutramine and placebo groups at the end point (systolic blood pressure: −2.1 vs −2.1 mm Hg; diastolic blood pressure: −0.1 vs −1.1 mm Hg; pulse rate: −0.2 vs −1.8 bpm). In both treatment groups, these reductions in vital signs were greater at the end point when BMI reduction was ≥5% compared with <5%.

CONCLUSIONS. Sibutramine may have some direct cardiovascular effects on obese adolescents. These cardiovascular effects may be balanced by a reduction in BMI, which, in adolescents, seems to be greater than that observed in adults.
OBESITY IS AN increasingly important medical problem among children and adolescents. Data from the National Health and Nutrition Examination Surveys show that the combined prevalence of overweight and at-risk-for-overweight alone among US children and adolescents has more than doubled since the early 1970s, whereas the prevalence of overweight has increased fourfold. Today, 16.1% of adolescents aged 12 to 19 years are overweight (≥95th percentile for age- and gender-specific measures for BMI).1,2 This epidemic has led to concern regarding the management of overweight and its complications. Although prevention is the ideal strategy, even with appropriate preventive approaches, it is likely that many children will become overweight and require treatment to prevent the long-term consequences of adult obesity, such as cardiovascular morbidity and mortality.3

In adults, excess body fat is associated with an increased risk for developing type 2 diabetes, heart disease, and a variety of other obesity-related conditions.4 Many of these diseases, previously considered diseases of adults, are now affecting children. In particular, overweight children and adolescents are known to suffer from hypertension, dyslipidemia, left ventricular hypertrophy, sleep apnea, and social and psychological problems. There has also been a dramatic increase in the incidence of type 2 diabetes in the pediatric population, for which the most important risk factor is obesity.5-8

The first-line therapy for overweight and obesity in both adults and adolescents is reduced-calorie diet, increased physical activity, and behavior modification.4,7,9 In overweight adolescents, comprehensive behavioral approaches have shown promise in short-term studies. Research is mixed with regard to whether adolescents benefit from family-based or individual behavioral programs.10,11 If family-based programs are initiated when the child is relatively young (6–12 years), the effects can persist into young adulthood. This argues, therefore, for starting weight management at an early age. Early treatment of overweight is also supported by a number of behavioral and metabolic factors, including a shorter history of the habits that lead to overweight. Because children are still growing and have a larger relative increase in height than weight, they are able to decrease their percent overweight.

Nevertheless, despite widespread and long-standing appreciation of the importance of diet and exercise to reduce childhood overweight, the evidence suggests that weight-loss programs for adolescents may be unsuccessful.12,13 In adults, such approaches are often supplemented by adjunctive drug therapy.4 There are no pharmacologic agents currently indicated for the management of pediatric overweight; however, in the United States, the approved labeling for orlistat has been modified to add pediatric use information.14 The serotonin and norepinephrine reuptake inhibitor sibutramine is approved for the management of adult obesity. It has been shown to produce dose-related weight loss and long-term weight maintenance with improvements in obesity-associated comorbidities for ≥2 years.15-17 However, in clinical trials, treatment with sibutramine in some patients is associated with small mean increases in systolic and diastolic blood pressure (SBP and DBP) and pulse rate (PR) relative to placebo.18,19 Given the potential consequences of long-term elevations in blood pressure, and because both the risks and benefits associated with pharmacotherapy cannot be assumed to be the same in this younger patient population as for adults, a full understanding of the benefit/risk profile of sibutramine in adolescents is important. The present analysis considers the cardiovascular risks of sibutramine in conjunction with behavioral therapy (BT) in a randomized, double-blind, placebo-controlled trial in obese adolescents.20

METHODS

This was a 12-month, randomized, double-blind, placebo-controlled study designed to evaluate the efficacy and safety of sibutramine in addition to BT in obese adolescents 12 to 16 years of age. The study was sponsored by Abbott Laboratories (previously Knoll Pharmaceuticals) and conducted at 33 centers in the United States from July 2000 to February 2002. Individual institutional review boards reviewed and approved the study. Written informed consent was obtained from the parent or legal guardian of each subject, and assent was obtained from the adolescent. Abbott Laboratories conducted data management and analysis in association with the lead investigators who contributed significantly to the development of the study protocol.

Adolescents aged 12 to 16 years in good general health with a BMI (calculated as weight in kilograms divided by the square of the height in meters) not less than a lower limit of ≥2 units above the US weighted mean for the 95th percentile based on age and gender21 and ≤44 kg/m² were enrolled in the study. Contraindications included cardiovascular disease (including arrhythmias), type 1 or type 2 diabetes mellitus, major psychiatric disorders, pregnancy, use of medications promoting weight loss or weight gain or those contraindicated with the use of sibutramine,22 or cigarette smoking. Candidates with SBP >130 mm Hg, DBP >85 mm Hg, or a PR >95 bpm were excluded, but hypertensive subjects stable on therapy were permitted.

Study Procedures

After an initial screening visit, eligible subjects were randomly assigned in a 3:1 ratio to receive either single daily doses of sibutramine 10 mg or placebo (Fig 1). Randomization was stratified by center and by low (≤37 kg/m²) and high (>37 kg/m²) baseline BMI. At month 6, all of the subjects who had not lost >10% of their initial
BMI were up-titrated in a blinded fashion to 15 mg of sibutramine or placebo. Study drugs were manufactured and supplied by Abbott Laboratories; identical placebo capsules were dispensed in the same way.

Subjects whose vital signs increased at a single visit to >150 mm Hg (SBP), and/or 95 mm Hg (DBP), and/or 110 bpm (PR) or increased by 20 mm Hg (SBP), and/or 15 mm Hg (DBP), and/or 20 bpm (PR) from baseline were either discontinued from study medication or closely monitored until their blood pressure returned to within acceptable limits. This decision was made at the discretion of the study physician. There was a protocol-driven algorithm for repeating visits and/or withdrawing subjects whose vital signs did not return to protocol-defined limits. There was no provision, however, for dose reduction from 15 to 10 mg in the event of intolerance. Throughout the study, all of the participants received site-specific instruction in lifestyle behavior modification.

Subjects were seen weekly until week 2, biweekly until week 12, and then monthly until study completion. At each scheduled visit, subjects were assessed for receipt of behavior modification instruction, and medication adherence was evaluated by capsule count.

Full physical examinations and focused cardiovascular examinations were also performed. Blood pressure was determined at all of the visits manually with the subject seated using an appropriately sized cuff and a calibrated mercury or aneroid sphygmomanometer. The onsets of the first and fifth Korotkoff phases were used to determine SBP and DBP, respectively. PRs were measured by palpation of the radial or brachial artery for ≥30 seconds. All of the readings were taken at approximately the same time of day. Three readings were taken at 2-minute intervals and averaged. Adverse events were assessed and recorded at each visit.

Data Analysis

All of the data are presented according to randomly assigned treatment group. Demographic and subject disposition data were summarized. All of the subjects receiving ≥1 dose of study medication were included in the safety summary of adverse events; 8 subjects recorded no postbaseline data and were excluded from all of the other analyses. The end point was defined as the latest postbaseline result. Statistical significance was to be determined by reference to the 5% level. The primary measure of efficacy was absolute change from baseline to the end point in BMI. Additional secondary efficacy variables that included anthropometric, glycemic, and lipids parameters were also determined and have been reported elsewhere.

Adverse events were summarized by using Coding Symbols for Thesaurus of Adverse Reaction Terms V. Fisher’s exact test was used to test for differences between treatment groups.

Vital signs data were summarized at each visit. The change from baseline to the end point in SBP, DBP, and PR was analyzed using analysis of covariance with a factor for treatment group and baseline value as a covariate. In an exploratory analysis, the same analyses were performed for the patient subgroups achieving <5%, ≥5%, and ≥10% reduction in BMI at the end point.

The number and percentage of subjects, with either values above absolute thresholds or high changes from baseline, using the protocol-specified criteria in Table 1 were summarized (“vital sign outlier events”). Adverse events associated with these outlier events were reported.

To assess the persistence of elevations in vital signs after baseline, the proportions of subjects with selected elevations at any visit, at ≥3 visits, and at ≥3 consecutively.
TABLE 1  Protocol-Defined Vital Sign Outlier-Event Criteria

<table>
<thead>
<tr>
<th>Outlier Variable</th>
<th>Single Visit</th>
<th>Three Consecutive Visits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absolute Threshold</td>
<td>Change From Baseline</td>
</tr>
<tr>
<td>SBP, mm Hg</td>
<td>&gt;150</td>
<td>&gt;20</td>
</tr>
<tr>
<td>DBP, mm Hg</td>
<td>&gt;95</td>
<td>&gt;15</td>
</tr>
<tr>
<td>PR, bpm</td>
<td>&gt;110</td>
<td>&gt;20</td>
</tr>
</tbody>
</table>

NA indicates not applicable.

tive visits were summarized. To compare the observed blood pressure data to expected values in this adolescent population, subjects were classified in an exploratory analysis as normal (<90th percentile), prehypertensive (90th to <95th percentile or, if blood pressure exceeded 120/80 mm Hg, even if <90th percentile up to <95th percentile), stage 1 hypertensive (95th to 99th percentile plus 5 mm Hg), or stage 2 hypertensive (>99th percentile plus 5 mm Hg), adjusting for age, height, and gender at baseline and the end point. The highest follow-up rating for each subject was also obtained. All of the statistical analyses were performed by using SAS 6.12 (SAS Institute, Cary, NC).

RESULTS

A total of 498 subjects were randomly assigned to 10 mg of sibutramine or placebo. The study population was 56.6% white, 21.1% black, and 15.7% Hispanic/Mexican American. At baseline, the treatment groups were well balanced (Table 2), and there were no important differences between treatment groups with respect to SBP or DBP, although seated PR was higher in the sibutramine treatment group (75.2 bpm) compared with the placebo treatment group (75.2 bpm).

Overall, 72% of the enrolled subjects completed the study: 76% (281 of 368) of the subjects in the sibutramine group and 62% (80 of 130) in the placebo group (P = .001). Postbaseline data were not recorded in 8 subjects (5 taking sibutramine and 3 taking placebo).

Mean exposure to study drug was 294 and 254 days in the sibutramine and placebo groups, respectively. Mean compliance was high (sibutramine: 89.1%; placebo: 83.9%). Overall, 47.9% subjects (174 of 363) in the sibutramine treatment group required a dose increase to 15 mg at month 6; mean exposure to 15 mg of sibutramine was 159 days.

Efficacy

Treatment with sibutramine plus BT resulted in a statistically significant reduction in mean (SE) BMI from baseline to the end point in the intention-to-treat population of −2.9 kg/m² (0.15 kg/m²) compared with −0.3 kg/m² (0.24 kg/m²) for placebo plus BT with a mean treatment difference of 2.6 kg/m² (SE: 0.27 kg/m²; 95% confidence interval: 2.0–3.1; P < .001), in favor of sibutramine.20 A BMI reduction of ≥5% and ≥10% occurred in 62.3% and 38.8% of subjects treated with sibutramine compared with 18.1% and 5.5% of subjects treated with placebo, respectively (odds ratio: 10.1 for ≥5% and 14.2 for ≥10% reduction; P < .001 for each).

Adverse Events

The incidence of adverse events reported was similar between subjects in the sibutramine and placebo treatment groups. Only 1 event, tachycardia, resulted in a significant difference between treatment groups (13% [46 of 368] sibutramine vs 6% [8 of 130] placebo; P = .049). Overall, adverse events led to 6% (23 of 368) withdrawals in the sibutramine group and 5% (7 of 130) in the placebo group (P = .832). Tachycardia and hypertension were the adverse events most commonly resulting in premature discontinuation of study drug. Of the 2% (9 of 368 and 2 of 130) of subjects discontinued for tachycardia in each treatment group, all had also been

TABLE 2  Baseline Characteristics for Subjects Assigned to BT Plus Sibutramine or Placebo

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Sibutramine Plus BT</th>
<th>Placebo Plus BT</th>
<th>Total Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 368)</td>
<td>(n = 130)</td>
<td>(N = 498)</td>
</tr>
<tr>
<td>Age, mean (SD), y</td>
<td>13.7 (1.3)</td>
<td>13.6 (1.3)</td>
<td>13.7 (1.3)</td>
</tr>
<tr>
<td>Weight, mean (SD), kg</td>
<td>97.9 (14.7)</td>
<td>97.8 (14.6)</td>
<td>97.9 (14.6)</td>
</tr>
<tr>
<td>Height, mean (SD), cm</td>
<td>164.3 (7.6)</td>
<td>164.9 (7.9)</td>
<td>164.7 (7.7)</td>
</tr>
<tr>
<td>BMI, mean (SD) kg/m²</td>
<td>36.1 (3.8)</td>
<td>35.9 (4.1)</td>
<td>36.1 (3.9)</td>
</tr>
<tr>
<td>Waist circumference, mean (SD), cm</td>
<td>105.8 (10.5)</td>
<td>106.3 (9.9)</td>
<td>106.0 (10.3)</td>
</tr>
<tr>
<td>No. (%) female subjects</td>
<td>242 (65.8)</td>
<td>80 (61.5)</td>
<td>322 (64.7)</td>
</tr>
<tr>
<td>No. (%) per race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>206 (56.0)</td>
<td>76 (58.5)</td>
<td>282 (56.6)</td>
</tr>
<tr>
<td>Black</td>
<td>80 (21.7)</td>
<td>25 (19.2)</td>
<td>105 (21.1)</td>
</tr>
<tr>
<td>Hispanic/Mexican American</td>
<td>60 (16.3)</td>
<td>18 (13.8)</td>
<td>78 (15.7)</td>
</tr>
<tr>
<td>Other</td>
<td>22 (6.0)</td>
<td>11 (8.5)</td>
<td>33 (6.6)</td>
</tr>
<tr>
<td>SBP (SD), mm Hg</td>
<td>113.3 (9.1)</td>
<td>113.2 (9.5)</td>
<td>113.3 (9.2)</td>
</tr>
<tr>
<td>DBP (SD), mm Hg</td>
<td>69.0 (7.6)</td>
<td>69.3 (7.4)</td>
<td>69.1 (7.6)</td>
</tr>
<tr>
<td>PR (SD), bpm</td>
<td>77.2 (8.7)</td>
<td>75.2 (8.9)</td>
<td>76.7 (8.8)</td>
</tr>
</tbody>
</table>

a n = 367.

b n = 497.
identified as PR outliers. Only 1% (5 of 368) of subjects in the sibutramine group (and none in the placebo group) were discontinued for hypertension; 4 of these 5 subjects were identified as blood pressure outliers.

Changes in Vital Signs

There were no violations of the assumptions of normality for the changes from baseline to the end point in blood pressure or PR data. In the intention-to-treat analysis, after adjusting for baseline values, the mean change in vital signs from baseline to the end point for SBP was $-2.1$ and $-2.1$ mm Hg for the sibutramine and placebo groups, respectively; for DBP, $-0.1$ and $-1.1$ mm Hg, respectively; and for PR, $-0.2$ and $-1.8$ bpm, respectively. The differences between treatment groups for the changes in blood pressure were not statistically significant ($P = .988$ and .136, respectively). For PR, the difference between treatment groups was $P = .055$. The mean changes in vital signs over time are displayed in Figs 2 through 4. Overall, mean changes from baseline to each visit were small and not clinically significant. For sibutramine, SBP ranged from $-2.4$ to $0.3$ mm Hg and DBP $-0.9$ to $1.5$ mm Hg and for placebo, SBP $-3.3$ to $-0.1$ mm Hg and DBP $-2.9$ to $-0.4$ mm Hg. For PR, mean changes ranged from $-0.9$ to $2.6$ bpm for sibutramine and $-1.4$ to $1.2$ bpm for placebo.

In subjects who completed the study, the mean absolute changes in the last 6 months of the study did not increase with continued drug exposure in either group. Subjects who were up-titrated to $15$ mg of sibutramine at month 6 did not demonstrate the same reduction in SBP, DBP, and PR as did the placebo and sibutramine 10-mg groups (data not shown). In an exploratory analysis, the absolute changes from baseline to the end point were analyzed for subjects by change in BMI (Table 3).

There were no statistically significant differences in the mean changes to the end point for any variable between subjects taking sibutramine and placebo who achieved BMI reductions of $\geq5\%$ or $\geq10\%$. The differences in mean changes in vital signs were statistically significant ($P \leq .02$) in subjects taking sibutramine who achieved a BMI reduction of $<5\%$ compared with placebo subjects who achieved a BMI reduction of $<5\%$. In both treatment groups, reductions to the end point in SBP, DBP, and PR were greater for subjects with $\geq5\%$ reduction in BMI at the end point compared with subjects with $<5\%$ change.

The overall incidence of subjects who ever recorded a vital sign outlier event was $32\%$ ($117$ of $363$) in the sibutramine group and $17\%$ ($21$ of $127$) in the placebo group ($P = .001$ for the difference). The incidence for blood pressure outlier events was similar in both treatment groups (SBP, sibutramine versus placebo: $5\%$ $[19$ of $363]$ vs $4\%$ $[5$ of $127]$; DBP, sibutramine versus placebo: $12\%$ $[43$ of $363]$ vs $8\%$ $[10$ of $127]$; respectively [P = .218]). There was a statistically significant difference between treatment groups in the incidence of PR outlier events (sibutramine versus placebo: $21\%$ $[75$ of $363]$ vs $6\%$ $[8$ of $127]$; respectively [P < .001]; Table 4).

Vital signs outlier events were reported more frequently during the first half of the study for both treatment groups. The number of subjects reporting $\geq1$ event before month 6 was $26\%$ ($95$ of $368$) in the sibutramine group and $14\%$ ($18$ of $130$) in the placebo group, compared with the number reporting $\geq1$ event after month 6, which was $14\%$ ($43$ of $310$) in the sibutramine group and $4\%$ ($4$ of $91$) in the placebo group.

Of the 138 subjects who experienced a vital sign outlier event during the treatment period, 51 had $\geq1$
vital sign–related adverse event (38.5% sibutramine [45 of 117] and 28.6% placebo [6 of 21]; eg, hypertension or tachycardia). The majority of these events were reported as mild in severity, and none was considered severe. There were no reports of serious adverse events related to vital signs in the subjects who experienced an outlier event, and the proportion of subjects who prematurely discontinued the study drug because of adverse events that were related to vital sign outlier events was similar (9.4% sibutramine [11 of 117] versus 9.5% placebo [2 of 21]).

Maintenance of Vital Sign Changes
At multiple and consecutive study visits, the proportion of subjects who had elevated SBP and DBP values above the absolute threshold or from baseline was small and similar between groups (Table 5). This was also the case for absolute PR increases, although the percentages were larger but not significant in the sibutramine group compared with the placebo group. Exploratory analyses found no correlation between maximum SBP or DBP and associated PR or between maximum PR and associated SBP or DBP.

Stratification According to Baseline Blood Pressure
In an exploratory analysis, subjects were stratified according to their baseline blood pressure.20 The analysis in Table 6 shows baseline blood pressure classification and the maximum follow-up classification during the treatment period. At baseline, blood pressure classification for all of the subjects was normal, prehypertensive, or stage 1; no subjects were classified with stage 2 blood pressure elevation. Only 7 (SBP = 5 and DBP = 2) of the 363 subjects in the sibutramine group and 1 (SBP) of the
127 subjects in the placebo group recorded shifts in their values to stage 2 hypertension during the treatment period, and only 1 subject (10 mg of sibutramine) remained at stage 2 after repeat testing. This subject presented with increased SBP on day 83, and after repeat blood pressure measurements, study medication was permanently withdrawn on day 166, and the subject was discontinued from the study. The event resolved 15 days after sibutramine was stopped. No subject in either group was classified with stage 2 hypertension at the end point.

**DISCUSSION**

The results of this study show that obese adolescents can achieve a significantly greater reduction in BMI when sibutramine, as compared with placebo, is used in combination with BT. The anthropometric changes in obese adolescents treated with sibutramine are similar to the results observed previously in obese adults.\(^1\)\(^,\)\(^3\)\(^,\)\(^5\)\(^-\)\(^17\) Importantly, the study suggests that sibutramine is well tolerated in this adolescent population. The occurrence of adverse events reported among obese adolescents was also comparable to the results from clinical trials of sibutramine treatment in obese adults.\(^1\)\(^,\)\(^3\)\(^-\)\(^17\) The study had an overall completion rate of 72%, which is high for a 1-year weight-loss study. Overall, 6% of subjects in the sibutramine group withdrew from this study compared with 5% in the placebo group \((P = .832)\).

In previous clinical studies in obese adults, sibutramine treatment has been associated with small mean

### TABLE 3  
Mean Change in Vital Signs for \(<5\%\), \(\geq5\%\), and \(\geq10\%\) BMI Responders From Baseline to the End Point

<table>
<thead>
<tr>
<th>Treatment Group</th>
<th>(&lt;5%) ↓ BMI</th>
<th>(\geq5%) ↓ BMI</th>
<th>(\geq10%) ↓ BMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Mean BL (SE)</td>
<td>Mean Change (SE)</td>
<td>N</td>
</tr>
<tr>
<td>SBP, mm Hg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sibutramine</td>
<td>137</td>
<td>114.2 (0.8)</td>
<td>0.0 (0.6)*</td>
</tr>
<tr>
<td>Placebo</td>
<td>104</td>
<td>113.0 (0.9)</td>
<td>−1.9 (0.6)</td>
</tr>
<tr>
<td>DBP, mm Hg</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sibutramine</td>
<td>137</td>
<td>68.8 (0.7)</td>
<td>1.8 (0.6)*</td>
</tr>
<tr>
<td>Placebo</td>
<td>104</td>
<td>69.2 (0.8)</td>
<td>−0.5 (0.6)</td>
</tr>
<tr>
<td>PR, bpm</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sibutramine</td>
<td>137</td>
<td>77.4 (0.8)</td>
<td>0.8 (0.7)*</td>
</tr>
<tr>
<td>Placebo</td>
<td>104</td>
<td>75.7 (0.9)</td>
<td>−1.7 (0.8)</td>
</tr>
</tbody>
</table>

*BL indicates baseline mean. Means and SEs at baseline were calculated by using analysis-of-variance models with a factor for treatment group. The analysis-of-covariance models for change also included baseline value.

* SBP: \(P = .033\) for treatment difference; PR: \(P = .020\) for treatment difference.

*b DBP: \(P = .007\) for treatment difference.

### TABLE 4  
Vital Sign Observations Meeting Outlier-Event Criteria

<table>
<thead>
<tr>
<th>Elevation Category</th>
<th>Sibutramine Plus BT ((N = 363), n (%))</th>
<th>Placebo Plus BT ((N = 127), n (%))</th>
<th>Total ((N = 490), n (%))</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;150 mm Hg</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>&gt;20 mm Hg greater than BL</td>
<td>18 (5)</td>
<td>5 (4)</td>
<td>23 (5)</td>
</tr>
<tr>
<td>&gt;150 and &gt;20 mm Hg greater than BL</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>&gt;15 but ≤20 mm Hg greater than BL on 3 consecutive visits</td>
<td>1 (0)</td>
<td>0 (0)</td>
<td>1 (0)</td>
</tr>
<tr>
<td>DBP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;95 mm Hg</td>
<td>3 (1)</td>
<td>0 (0)</td>
<td>3 (1)</td>
</tr>
<tr>
<td>&gt;15 mm Hg greater than BL</td>
<td>38 (10)</td>
<td>10 (8)</td>
<td>48 (10)</td>
</tr>
<tr>
<td>&gt;95 and &gt;15 mm Hg greater than BL</td>
<td>2 (1)</td>
<td>0 (0)</td>
<td>2 (0)</td>
</tr>
<tr>
<td>&gt;10 but ≤15 mm Hg greater than BL on 3 consecutive visits</td>
<td>8 (2)</td>
<td>0 (0)</td>
<td>8 (2)</td>
</tr>
<tr>
<td>PR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;110 bpm</td>
<td>10 (3)</td>
<td>0 (0)</td>
<td>10 (2)</td>
</tr>
<tr>
<td>&gt;20 bpm greater than BL</td>
<td>72 (20)</td>
<td>8 (6)</td>
<td>80 (16)</td>
</tr>
<tr>
<td>&gt;110 and &gt;20 bpm greater than BL</td>
<td>9 (2)</td>
<td>0 (0)</td>
<td>9 (2)</td>
</tr>
<tr>
<td>&gt;105 but ≤110 or &gt;15 but ≤20 bpm greater than BL on 3 consecutive visits</td>
<td>6 (2)</td>
<td>1 (1)</td>
<td>7 (1)</td>
</tr>
</tbody>
</table>

*BL indicates baseline.

* Categories are not mutually exclusive.

* Subjects in this category are also included in both single-elevation categories for the respective vital sign.
increases in blood pressure and PR relative to placebo in some obese subjects. Published clinical studies in overweight adolescents are limited; only one has shown effects similar to adults. Although the long-term consequences of untreated or undertreated hypertension in adults are well documented, they are less clear in children and adolescents. In addition, no data are available on the long-term effects of antihypertensive drugs on the cardiovascular effects of sibutramine in this population.

The present study resulted in statistically significant and clinically important improvements in BMI. A mean absolute change in BMI of \(-2.9 \text{ kg/m}^2\) (\(-8.2\%\)) for the sibutramine group compared with \(-0.3 \text{ kg/m}^2\) (\(-0.8\%\); \(P < .001\) for both) for the placebo group was evident at the end point. The odds ratio for achieving \(\geq 5\%\) BMI reduction at the end point with sibutramine treatment compared with placebo was 10.1 (\(P < .001\)). Furthermore, there were no statistically significant differences in blood pressure between obese adolescents treated with sibutramine and those given placebo. Both treatment groups showed small mean decreases in SBP, DBP, and PR. In those subjects with \(\geq 5\%\) reduction in BMI at the end point, these decreases in blood pressure and PR were greater compared with subjects with \(<5\%\) reduction in BMI. The decrease in mean blood pressure in subjects treated with sibutramine was less than that in subjects who achieved an equivalent change in BMI by BT alone, but it is important to consider the effectiveness of BT to induce a decrease in BMI. Only 38.8\% of the adolescents treated with BT alone achieved a reduction in their BMI of \(>5\%\) compared with 62.3\% of those treated with sibutramine plus BT. In subjects with \(>10\%\) weight loss, the difference was even more marked: 5.5\% of subjects treated with BT alone compared with 18.1\% when sibutramine was added to BT.

These findings are similar to those seen in adult subjects who achieve clinically meaningful weight loss (ie, \(\geq 5\%\)) when treated with sibutramine and lifestyle modification. Exploratory analyses were performed to assess the magnitude and persistence of vital sign changes between the treatment groups, while accounting for the substantial intrasubject variability in blood pressure and PR observed in adolescents. The incidence of vital signs outlier events was higher in subjects randomly assigned to sibutramine because of a statistically significant difference in PR outliers between groups. However, the increases in vital sign measurements did not persist over time and, therefore, are not likely to be clinically important. The data are reassuring, because they included values recorded from the first few weeks of therapy with sibutramine when weight loss was at a minimum and its potential beneficial effects on blood pressure would be absent. In addition, there was no correlation between PR and SBP or PR and DBP. The primary end point of this study was change in BMI, and although the study was not specifically powered to assess changes in blood pressure and vital signs, a sample size of 100 placebo subjects was considered adequate to assess the anticipated treatment-group differences.

It is not clear why blood pressure and heart rate respond differently to treatment with sibutramine in adolescents compared with adults. It may be that autonomic balance differs in adolescents, favoring the parasympathetic system, because vagal tone is known to be higher in adolescents. If this is the case, the sympathetic activation characteristically associated with sibutramine may be counteracted.

Sibutramine is a serotonin and norepinephrine reuptake inhibitor, which is generally assumed to act centrally. Recent research suggests that the cardiovascular effects of sibutramine may result from a complex inter-

### Table 5: Persistence of Increases in Vital Signs

<table>
<thead>
<tr>
<th>Vital Sign Variable and Measurement</th>
<th>Sibutramine Plus BT (n = 363), %</th>
<th>Placebo Plus BT (n = 127), %</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\geq 130 \text{ mm Hg})</td>
<td>10.7</td>
<td>10.2</td>
</tr>
<tr>
<td>At (\geq 1) visit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At (\geq 3) visits</td>
<td>3.0</td>
<td>3.9</td>
</tr>
<tr>
<td>At (\geq 3) consecutive visits</td>
<td>1.4</td>
<td>0.8</td>
</tr>
<tr>
<td>Increase (\geq 10 \text{ mm Hg}) from baseline</td>
<td>36.1</td>
<td>25.2</td>
</tr>
<tr>
<td>At (\geq 1) visit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At (\geq 3) visits</td>
<td>14.9</td>
<td>11.8</td>
</tr>
<tr>
<td>At (\geq 3) consecutive visits</td>
<td>6.9</td>
<td>3.9</td>
</tr>
<tr>
<td>Increase (\geq 20 \text{ mm Hg}) from baseline</td>
<td>6.6</td>
<td>4.7</td>
</tr>
<tr>
<td>At (\geq 1) visit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At (\geq 3) visits</td>
<td>0.8</td>
<td>0.0</td>
</tr>
<tr>
<td>At (\geq 3) consecutive visits</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>DBP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\geq 85 \text{ mm Hg})</td>
<td>1.4</td>
<td>0.0</td>
</tr>
<tr>
<td>At (\geq 1) visit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At (\geq 3) visits</td>
<td>0.3</td>
<td>0.0</td>
</tr>
<tr>
<td>Increase (\geq 10 \text{ mm Hg}) from baseline</td>
<td>13.2</td>
<td>12.6</td>
</tr>
<tr>
<td>At (\geq 1) visit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At (\geq 3) visits</td>
<td>4.1</td>
<td>3.1</td>
</tr>
<tr>
<td>At (\geq 3) consecutive visits</td>
<td>1.9</td>
<td>1.6</td>
</tr>
<tr>
<td>PR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(&gt;95 \text{ bpm})</td>
<td>23.4</td>
<td>9.4</td>
</tr>
<tr>
<td>At (\geq 1) visit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At (\geq 3) visits</td>
<td>5.5</td>
<td>0.8</td>
</tr>
<tr>
<td>At (\geq 3) consecutive visits</td>
<td>3.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Increase (\geq 10 \text{ bpm}) from baseline</td>
<td>53.2</td>
<td>43.3</td>
</tr>
<tr>
<td>At (\geq 1) visit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At (\geq 3) visits</td>
<td>28.9</td>
<td>22.0</td>
</tr>
<tr>
<td>At (\geq 3) consecutive visits</td>
<td>14.0</td>
<td>11.0</td>
</tr>
<tr>
<td>Increase (\geq 20 \text{ bpm}) from baseline</td>
<td>17.9</td>
<td>8.7</td>
</tr>
<tr>
<td>At (\geq 1) visit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At (\geq 3) visits</td>
<td>5.0</td>
<td>3.1</td>
</tr>
<tr>
<td>At (\geq 3) consecutive visits</td>
<td>1.7</td>
<td>1.6</td>
</tr>
</tbody>
</table>

*Categories are not mutually exclusive.*
action of peripheral and central nervous system effects. For example, mechanisms that drive the peripheral effects of sibutramine dominate and promote resting blood pressure increases in healthy young subjects with low sympathetic activity. It is also proposed that the central clonidine-like inhibitory effects of sibutramine become apparent when sympathetic nervous system activity is acutely increased, attenuating the pressor response to sympathetic stimuli. This may explain the blood pressure reduction observed in obese subjects. The results of the present study suggest a dominant role for the central clonidine-like action. This highlights the need for additional work to elucidate the precise mechanisms by which sibutramine and possibly weight loss are acting to alter blood pressure in overweight adolescents. We should be careful, however, not to overstate the apparent differences between adults and adolescents as shown in this study, and in the absence of direct comparative studies, these findings should be considered as suggestive only. In addition, it must be appreciated that these study results cannot be generalized to include overweight adolescents with hypertension or tachycardia, because subjects with these conditions were excluded from entry. At this time, the Food and Drug Administration has not approved sibutramine use in adolescents who are <16 years of age.

Increases in obesity-related morbidity and mortality carry over to adulthood from adolescence. In the Harvard Growth Study, adolescents at risk of overweight (BMI >75th percentile) were significantly more prone to adult morbidity and mortality from coronary heart disease and mortality from all causes, independent of adult weight status. In the Boyd Orr cohort, all-cause and cardiovascular mortality was increased among adults who were in the ≥75th BMI percentile as children and adolescents. The adverse effects of adolescent overweight on cardiovascular risk factors have been evaluated in the Bogalusa Heart Study. Adverse levels of lipids, fasting insulin, and blood pressure were more common in adolescents with a BMI >95th percentile compared with those with a BMI <85th percentile. Clustering of ≥2 adverse risk factors occurred in 50% of the high-BMI group. Follow-up investigations revealed that elevated BMI and adverse risk factor clustering persisted from adolescence into young adulthood. Furthermore, increased BMI was associated with asymptomatic atherosclerosis in the aorta and coronary arteries.

Nearly 30% of adolescents with BMI >95th percentile are estimated to have metabolic syndrome according to the National Cholesterol Program compared with ~7% with a BMI between the 85th and <95th percentile. In adolescents with a BMI <85th percentile, the rate is only 0.1%. There are no definitive guidelines for assessing clinically significant changes in weight status and related comorbidities in adolescents and no long-term outcome studies that demonstrate the positive effects of weight reduction in this population. Substantial evidence from the treatment of overweight and obese adults corroborates that moderate weight loss (≥5% to 10%) is associated with improvements in obesity-related comorbidities, delay of onset of type 2 diabetes, and reduced mortality because of obesity. It is reasonable, therefore, to propose that the positive effects of weight loss in obese adults are applicable to weight loss in overweight adolescents. Thus, the observed effects of sibutramine treatment on surrogate markers for the development of cardiovascular disease, type 2 diabetes, and increased mortality because of obesity in

<table>
<thead>
<tr>
<th>Variable</th>
<th>Overall, n</th>
<th>Maximum Follow-up Classification, n</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal Prehypertension Stage 1 Stage 2</td>
<td></td>
</tr>
<tr>
<td>Baseline SBP classification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sibutramine plus BT</td>
<td>Normal 271 Prehypertension 142 Stage 1 95 33</td>
<td></td>
</tr>
<tr>
<td>Placebo plus BT</td>
<td>Normal 96 Prehypertension 63 Stage 1 29 4</td>
<td></td>
</tr>
<tr>
<td>Baseline DBP classification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sibutramine plus BT</td>
<td>Normal 325 Prehypertension 206 Stage 1 87 31</td>
<td></td>
</tr>
<tr>
<td>Placebo plus BT</td>
<td>Normal 110 Prehypertension 79 Stage 1 22 9</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 6 Maximum Follow-up Classification for SBP and DBP According to Baseline Classification

PEDIATRICS Volume 120, Number 1, July 2007 e155
obese adolescents should be considered clinically significant.

CONCLUSIONS
Overweight poses a major health problem for the pediatric population with limited treatment options. First-line therapy for overweight adolescents should be the same as that for overweight and obese adults, that is, a calorie-controlled diet, increased physical activity, and behavior modification. Pharmacotherapy should be used as adjunctive therapy only to lifestyle modification. In a pediatric population where the individuals are still actively growing, even small losses in BMI or, in fact, maintenance of BMI are important outcomes that will impact related comorbidities.

This study has demonstrated that sibutramine treatment effectively promotes weight loss in obese adolescents with concomitant improvements in blood pressure and heart rate. Sibutramine treatment seems to have minimal cardiovascular effects and to be well tolerated in this population.

ACKNOWLEDGMENTS
The study was supported by Abbott Laboratories.

The work of the staff in the participating centers and the assistance of all of the study subjects are gratefully acknowledged.

REFERENCES


Motor Profile of Children With Developmental Speech and Language Disorders

Chris Visscher, PhDa,b, Suzanne Houwen, MSCa,b, Erik J. A. Scherder, PhDa, Ben Moolenaar, BEdc, Esther Hartman, PhDa

aCenter for Human Movement Sciences and bUniversity Center for Sport, Movement, and Health, University Medical Center Groningen, University of Groningen, Groningen, Netherlands; cHanze Institute for Sports Studies, Hanze University Groningen, Groningen, Netherlands

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

ABSTRACT

OBJECTIVES. The purpose of this study was to investigate the motor profile of 125 children with developmental speech and language disorders and to test for differences, if any, in motor profile among subgroups of children with developmental speech and language disorders.

METHODS. The participants were 125 children with developmental speech and language disorders aged 6 to 9 years from 2 special schools for children with communication problems in the northern Netherlands. They were tested with the Movement Assessment Battery for Children. The children were classified by the schools’ speech and language therapists into 3 subgroups on the basis of language tests, oral motor tests, and clinical examinations: children with speech disorders (n = 14), language disorders (n = 46), or both (n = 65).

RESULTS. Compared with the norms of the Movement Assessment Battery for Children, children with developmental speech and language disorders performed significantly less well. Results showed that 51% of the children with developmental speech and language disorders had borderline or definite motor problems. Children with language disorders had significantly lower scores (ie, better performance) on the ball-skills subtest and the total test than children with speech disorders and children with both speech and language disorders. Furthermore, children with language disorders had significantly better performance on the balance subtest than children with both speech and language disorders.

CONCLUSIONS. The findings of this study support the idea that developmental speech and language disorders are frequently associated with motor problems and that the kind of developmental speech and language disorders affects motor performance differently. Speech and language disorders seem to have more impact on motor performance than only language disorders, and it seems that when speech production is affected, motor problems are more pronounced. The findings support the need to give early and more attention to the motor skills of children with developmental speech and language disorders in the educational and home setting, with special attention to children whose speech is affected.
DEVELOPMENTAL SPEECH AND LANGUAGE DISORDERS (DSLDs) are characterized by delays in speech and language development in the absence of mental or physical handicap, hearing loss, emotional disorder, or environmental deprivation. The clinical picture is quite varied; many children have speech as well as language disorders, others may have pure speech disorders or pure language disorders. The prevalence of DSLDs varies from 1.3% to 7.4%, depending on the definition used.

Although most attention has been paid to the communication profile of children with DSLDs, it has been shown that motor problems are not uncommon in this population. The co-occurrence of motor problems and DSLDs may be explained by both factors within the child, such as a genetic risk or neurologic deficits, and environmental factors, such as communication difficulties negatively influencing social acceptance and participation in play and sports activities.

The majority of studies concerning motor problems in children with DSLDs mainly focused on fine motor tasks. Studies indicate that these children are significantly slower than regular children on tasks that mainly challenge eye-hand coordination (ie, pegboard, threading beads, fastening buttons, and tapping). Of note is that motor problems seem to not be restricted to tasks involving time constraints. For gross motor skills, it has been observed that skills like stepping, running, stair climbing, muscle tone, standing on 1 leg, hopping on 1 leg, toe gait, heel gait, and skills that involve object control or locomotor activity of children with language problems were poor relative to regular children. Moreover, balancing on 1 leg proved to be 1 of the most discriminating measures between children with specific language impairment and regular children. In contrast, results of an early study found no difference between children with specific language impairment and regular children in duration of balance.

Quite clearly there is strong evidence of clinically significant overlap between DSLDs and motor problems; however, 2 things are of note. First, hardly any attention has been paid to ball skills of children with DSLDs, although these skills explicitly may challenge eye-hand coordination, depend on balance control, and importantly contribute to the child’s social interaction with peers. Because children with DSLDs may already have problems with social acceptance, because of their communication difficulties, inadequate ball skills may further restrict the child’s capacity to interact socially and physically with peers. Within this scope, it is noteworthy that recent epidemiologic studies emphasize the value of a social and physical active lifestyle, particularly when started early in life. One of the major effects of such a lifestyle is reducing the risk for cognitive impairment later in life.

Second, research examining the motor performance of subgroups of children with DSLDs is limited. Hill suggested that subgroups of children with DSLDs differ in their performance on fine motor tasks. Bishop addressed the issue of subtype-specific differences in relation to motor performance and found some interesting results. In 2 twin studies where 1 or both twins had speech/language impairment along with a control group of unaffected children, she found that children with combined speech and language impairments obtained poorer scores on a pegboard and tapping task than unaffected children. Furthermore, she concluded that the link between speech/language impairments and motor problems was stronger for speech than for language impairments. It is important to gain insight in the performance profile of subgroups of children with DSLDs, because this information may provide clues for effective intervention.

Specifically, this study had 2 aims. The first aim was to investigate the motor profile of children with DSLDs with respect to manual dexterity, ball skills, and balance. The second aim was to test for differences, if any, in motor profile among 3 subgroups of children with DSLDs: children with speech disorders, children with language disorders, and children with both speech and language disorders.

METHODS

Participants
A total of 125 children with DSLDs aged 6 to 9 years (93 boys and 32 girls; mean age: 7.4 years; SD: 1.1 years) were recruited from 2 special schools for children with communication problems in the northern Netherlands. Children with other impairments, like hearing impairments or physical impairments, were excluded from the study. No child had an intelligence quotient <80.

Subtests of the Dutch Language Test for Children (Taaltest voor Kinderen) have been used by the schools’ speech and language therapists to establish the diagnosis of the children. This test battery is composed of subtests that provide information about the child’s receptive and expressive language skills in the areas of morphology, syntax, and semantics. The word forms production test (woordvormen produktie test), the syntax production test (zinsbouw produktie test), and the vocabulary production test (woordenschat produktie test) were used to provide information about a child’s expressive language skills. The concealed meaning test (verzwegen betekenis test), the syntax choice test (zinsbouw keuze test), syntax evaluation test (zinsbouw beoordeling test), and the vocabulary choice test (woordenschat keuze test) were used to provide information about a child’s receptive language skills. Speech skills (oral motor function) were assessed with both oral motor tests and clinical examinations.

The children were considered to have language and/or speech problems when the speech and language...
therapists had noted a deviation of ≥1 SD compared with age standards. The schools’ speech and language therapists independently reviewed the available data to classify the children into 3 subgroups of DSLDs. The few discrepancies in classification were discussed until consensus was reached.

It was found that 46 children had language problems (34 boys and 12 girls; mean age: 7.7 years; SD: 1.1 years), of which 26 had expressive problems and 20 had mixed expressive-receptive problems. Sixty-five children had both speech and language problems (50 boys and 15 girls; mean age: 7.2 years; SD: 1.1 years), of which 43 had expressive language problems combined with oral motor problems and 22 had expressive-receptive language problems combined with oral motor problems. Fourteen children had oral motor problems (9 boys and 5 girls; mean age: 7.3 years; SD: 1.3 years). The 3 groups did not differ significantly on age (analysis of variance: $F_{2,122} = 2.88; P > .05$).

Informed consent to participate was obtained from the children’s parents. The procedures were in accordance with the ethical standards of the Faculty of Medical Sciences of the University of Groningen.

**Movement Assessment Battery for Children**

The Movement Assessment Battery for Children (ABC) is a test battery designed for diagnosis of delays or deficits in motor development. The Movement ABC consists of 4 age-related item sets (4–6, 7–8, 9–10, and 11–12 years). Each set is built up of 8 tasks, which are assessed under the following 3 subtests: manual dexterity, ball skills, and static and dynamic balance. Only the first 3 age-related item sets were used, because the study was performed with children from 6 to 9 years old.

Each item is scored on a scale from 0 to 5. Summing the item scores of the 3 subtests produces a profile of the child’s performance. The manual-dexterity subtest score varies from 0 to 15, the ball-skills subtest score from 0 to 10, and the static and dynamic balance subtest from 0 to 15. The 3 subtest scores can be summed to produce a total test score ranging between 0 and 40. High scores indicate poor motor performance. The 3 subtests scores and the total test score can be transformed into percentile scores that show the child’s level of performance in comparison with his or her peers. The range between the 100th and 16th percentile was regarded as “no motor problems,” 15th to 6th percentile as “borderline motor problems,” and the 5th percentile and below as “definite motor problems.”

The test has acceptable validity and reliability. Interrater reliability ranges from .70 to .89, and the test-retest reliability is .75. In an earlier pilot study with 59 children, it seemed that the test-retest reliability of the Movement ABC was .83 for children with communication problems. Van Waelvelde et al recently confirmed the concurrent validity of the total test score and the ball-catching item of the second age band of the Movement ABC.

**Procedure**

The Movement ABC was administered individually in a quiet room at the school of each participant. Depending on the age and performance level of the individual child, test duration ranged from 25 to 35 minutes. The same examiner conducted all of the testing. The examiner was blind to the subgroup diagnosis of the children with DSLDs. Movement ABC testing was conducted according to the manual of this test.

**Data Analysis**

The statistics were performed by using SPSS 11.0 (SPSS Inc, Chicago, IL). The difference between the group of children with DSLDs and the normative population with respect to the proportion of children with motor problems was assessed with the $\chi^2$ test. The nonparametric Kruskal-Wallis test was used to determine whether there were differences among median scores on the Movement ABC for the 3 subgroups of children with DSLDs with an $\alpha$ level set at .05. If a significant difference was found between the 3 groups, 2-group Mann-Whitney $U$ comparisons were made. Because type 1 error can be inflated by performing multiple statistical tests, the $\alpha$ level was corrected using the Bonferroni correction and set at .017 for each posthoc comparison (.05 divided by the 3 posthoc comparisons).

To assist in determining the meaningfulness of group effects, correlational effect size statistics for nonparametric data were calculated for each dependent variable. Effect size was calculated by dividing the $z$ score by the square root of the number of children contributing to the analyses. An effect size of $r = 0.10$ was defined as small, $r = 0.30$ as medium, and $r = 0.50$ as large.

**RESULTS**

**Motor Profile of Children With DSLDs**

On the basis of the test, the children were categorized as having definite motor problems, borderline motor problems, or no motor problems (see Table 1). Among the children with definite motor problems, the total test score of the Movement ABC ranged from 13.5 to 36.0.

<table>
<thead>
<tr>
<th>Test</th>
<th>Definite, $n$ (%)</th>
<th>Borderline, $n$ (%)</th>
<th>No Problems, $n$ (%)</th>
<th>$\chi^2$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual dexterity</td>
<td>25 (20)</td>
<td>16 (13)</td>
<td>84 (67)</td>
<td>61.89</td>
<td>.000</td>
</tr>
<tr>
<td>Ball skills</td>
<td>38 (31)</td>
<td>44 (35)</td>
<td>43 (34)</td>
<td>278.32</td>
<td>.000</td>
</tr>
<tr>
<td>Balance</td>
<td>23 (19)</td>
<td>19 (15)</td>
<td>83 (66)</td>
<td>53.36</td>
<td>.000</td>
</tr>
<tr>
<td>Total test</td>
<td>38 (30)</td>
<td>26 (21)</td>
<td>61 (49)</td>
<td>195.14</td>
<td>.000</td>
</tr>
</tbody>
</table>
The χ² test revealed that the proportions of children with DSLDs with borderline motor problems (21%) and definite motor problems (30%) differed significantly (P < .001) from the proportions that would have been expected in a normal population (10% and 5%, respectively). Furthermore, it can be seen that a similar picture was obtained for the 3 subtests (see Table 1). The χ² tests revealed that more children with DSLDs scored in the category “definite motor problems” and in the category “borderline motor problems” than a normal population for the manual-dexterity, ball-skill, and balance subtests (P < .001). This effect was most pronounced for ball skills.

Motor Profile of Subgroups of Children With DSLDs

Table 2 displays the means, medians, SDs, and ranges of performance for each subgroup of children with DSLDs for manual dexterity, ball skills, balance, and total test. Significant differences were found for the ball-skills (H₂ = 21.42; P < .001) and balance (H₂ = 7.39; P < .05) subtests and for the total test (H₂ = 13.85; P ≤ .001).

Mann-Whitney U tests with Bonferroni correction showed that children with language disorders scored significantly lower (ie, better performance) on the ball-skills subtest than children with speech disorders (z = −3.21; P ≤ .001; r = 0.42) and children with both speech and language disorders (z = −4.29; P < .001; r = 0.41). For the balance subtest, it was found that children with language disorders scored significantly lower than children with both speech and language disorders (z = −2.69; P < .01; r = 0.25). For the total test, the results showed that children with language disorders scores significantly lower than children with speech disorders (z = −2.52; P < .01; r = 0.33) and children with both speech and language disorders (z = −3.49; P ≤ .001; r = 0.33). These effect size statistics represent moderate-to-large effects, except for the balance subtest, where the effect was small-to-moderate. The other posthoc comparisons were nonsignificant with small effect-size statistics.

**DISCUSSION**

Our study showed that the overall motor performance of children with DSLDs was quantitatively impaired. The children with DSLDs demonstrated deficits in manual dexterity, ball skills, and static and dynamic balance as measured by the Movement ABC. The results revealed that 51% of the children with DSLDs had borderline or definite motor problems. Other studies that used the Movement ABC, or its predecessor the Test of Motor Impairment, in children with speech and language disorders found similar percentages, ranging from 40% to 90%. However, in these studies, no distinction was made between problems in the different subtests or skills. Furthermore, in many other studies often 1 aspect of motor performance, primarily fine motor skills, has been investigated. Striking was that in the present study motor problems in children with DSLDs seemed most pronounced for ball skills as measured by catching a moving object and throwing at a target. Wiznitzer et al had mentioned that problems could be expected in throwing and catching for children with developmental language disorder.

Why are motor problems in children with DSLDs so considerable? According to the atypical brain development framework, developmental variation in brain structures and functions leads to variation in abilities underscoring the interrelatedness of developmental disorders. Unfortunately, the atypical brain development framework does not address how specific areas of the brain explain particular abilities. In this case, it is interesting to discuss possible mechanisms underlying the present findings. Important brain structures involved in each of the 3 types of motor activity examined in our study are the basal ganglia. In catching a ball, the basal ganglia coordinate the duration of the reach and approach of the ball. Furthermore, the basal ganglia play a role in balance control and in some types of manual dexterity such as handwriting. Interestingly, a disturbance in language production and speech initiation could be viewed as a consequence of a disturbed function of the left basal ganglia. It is also interesting to discuss possible mechanisms underlying the present findings at the level of neural circuits. One of the structures of the basal ganglia is the caudate nucleus, which has a strong functional relationship with the prefrontal cortex. Damage to 1 of these regions may express itself in a decline in the control and execution of movements.
combined with cognitive deficits, among which are specific language disturbances. A similar clinical outcome emerges from damage to a neural circuit in which the prefrontal cortex closely cooperates with the cerebellum. A disturbance in this latter neural circuit might be responsible for impairments in timing precision found in children with specific language disorders or dyslexia.

Environmental factors, however, probably also play a role in the link between DSLDs and motor problems. Children with DSLDs experience communication difficulties, which may negatively influence social acceptance. Children who have lower levels of social acceptance are supposed to participate less in play with peers. As a consequence, a lack of practice of motor skills may occur, which may result in low levels of motor skills. When one considers that ball skills are a part of many play and sport activities, it seems obvious that less participation in play and sport activities may affect the adequate learning of these skills.

Regarding the effects of subgroup division on motor performance, we found that subgroups differed in motor performance: children with language disorders had better performance than children with speech disorders and children with both speech and language disorders on the ball-skills subtest and total test, and children with language disorders had better performance on the balance subtest than children with both speech and language disorders. Although the differences between children with language disorders and children with speech disorders were only significant for ball skills and the total test, it is worth mentioning that there was a tendency for the children with speech disorders to perform worse than children with language disorders and that their scores were going more in the direction of those obtained by the children with both speech and language disorders. In our study, only 14 children with speech disorders participated, so this is likely to account for the lack of significance. Thus, it seems that when speech production is affected, motor problems are more pronounced. The results are partly in line with the findings of Bishop. She found that children with both speech and language disorders aged 7 to 13 years showed the greatest motor problems as measured by a peg-moving and tapping task.

What could be the reason that children with speech disorders and children with both speech and language disorders had a worse performance on ball skills and the total test than children with only language disorders? Diversity in outcomes may be the result of the different neural circuits in which brain structures participate. Because of the marked segregation and specificity of inputs and outputs of brain structures, dysfunctions of slightly different parts of a specific brain structure may express themselves in different ways in subgroups of children with DSLDs. Furthermore, for children with both speech and language disorders, it is not inconceivable that having both disorders has a greater impact on social functioning and social behavior than only language disorders because of the more complicated communication.

Our study showed that motor problems were evident in a large proportion of children aged ≥6 years. Recently, it has been shown that co-occurrence of motor problems and language delays already exists at the age of 4 years. Because the role of verbal communication is less important in play and sports activities of very young children but becomes more and more important when children grow older, one might expect the motor problems to become even bigger. Because motor problems may have important consequences for both a child’s academic performance and a child’s ability to participate in play and sports activities, intervention at an early age is warranted. There is some evidence that interventions may lead to improvements in motor performance of children with speech and language problems. Information provided by the Movement ABC may be helpful in deciding to provide interventions or not. For children with a score below the 5th percentile, an intervention is imperative, but the mode and type may vary. For children with scores between the 5th and 15th percentile, the decision to intervene has to depend on the impact of the child’s motor problems on both daily life motor functioning, as well as other developmental areas, such as social functioning. Although a score below the 15th percentile on the Movement ABC is indicative of developmental coordination disorder (DCD), a child has to meet other criteria to be diagnosed as having DCD. This means that for the children in our study who scored below the 15th percentile on the Movement ABC, further investigation is warranted. Establishing a diagnosis of DCD was, however, not an aim of this study; therefore, we cannot assert with certainty that these children also have DCD.

A delimitation of this study was the small sample size of the children with only speech disorders and the missing information about the etiology of the disorders. This study, however, is the first onset to give a more comprehensive view of the motor profile of children with DSLDs and differences between subgroups. In future studies, more attention should be given to the motor profiles of children with speech, language, or both speech and language disorders.

CONCLUSIONS

Our results support the idea that DSLDs are frequently associated with motor problems and that the kind of DSLDs affects motor performance differently. Speech and language disorders seem to have more impact on motor performance than only language disorders, and it seems that when speech production is affected, motor problems are more pronounced. The results support the need to give early and more attention to motor devel-
development of children with DSLDs in the educational and home setting in addition to other developmental skills. Special attention should be given to children with both speech and language disorders.

ACKNOWLEDGMENTS
We thank M. C. B. van der Nagel for her assistance in this research project and all of the teachers, other staff, and children of the schools for their cooperation.

REFERENCES
OBJECTIVES. Studies of adults have shown that thrice-daily hydrocortisone dosing results in more physiologic cortisol profiles than twice-daily dosing. There are no data on thrice-daily dosing and only limited data on twice-daily dosing in children despite the possible adverse effects of glucocorticoid underreplacement or overreplacement.

METHODS. Using 24-hour cortisol and glucose profiles, along with computerized cognitive testing, our aim was to assess prescribed hydrocortisone regimens in children and adolescents with hypopituitarism.

RESULTS. Twenty patients with adrenocorticotrophic hormone deficiency participated. The hydrocortisone dosing regimen was thrice daily in 9 patients and twice daily in 11 patients (mean total daily dose: 8.3 ± 2.6 and 7.6 ± 2.1 mg/m² per day, respectively). Those on twice-daily dosing had more waking hours (between 8:00 AM and 8:00 PM) below the reference range than those on thrice-daily dosing (5.5 vs 2.1) and more daytime prolonged hypocortisolemia, defined as plasma cortisol level of <50 nmol/L for ≥4 hours (64% vs 0%). Morning doses >4 mg/m² caused larger postdose peaks than <4 mg/m² (151 vs 47 nmol/L, above the 97.5th percentile). However, there was no difference in the length of time taken to reach nadir below the 2.5th percentile (5.2 vs 4.8 hours). This was true for evening doses of >2.5 mg/m² and < 2.5 mg/m². No hypoglycemia or hyperglycemia was detected in association with low or high cortisol levels. On predose and postdose cognitive testing (34 paired tests), no significant change in reaction speed was detected (453.3 vs 438.8 milliseconds) or in subgroup analysis of those who had symptoms of lethargy, predose cortisol levels of <50 nmol/L, or prolonged hypocortisolemia.

CONCLUSIONS. Thrice-daily dosing resulted in less frequent and prolonged hypocortisolemia than twice-daily regimens, but we were unable to relate either regimen to acute clinical end points of glycemia, lethargy, or cognitive function.
The current management of adrenocorticotropic hormone (ACTH) deficiency involves daily hydrocortisone replacement plus increased doses during physiologic stress. The optimum hydrocortisone dose and replacement regimen remains unclear in ACTH deficiency, whether considered to be complete or partial. Estimates of daily cortisol production rates have been revised downward from 12 to 15 mg/m² per day to 6 to 8 mg/m² per day in recent times. Therefore, given the high oral bioavailability of hydrocortisone (≥95%), a dose of 6 to 8 mg/m² per day should be the aim of therapy, provided the patient has no hypoglycemia or symptoms of cortisol deficiency.

The short-term and medium-term risks of glucocorticoid underreplacement or overreplacement are particularly important in children and adolescents. In addition, longer-term health into adulthood needs to be considered. Adults with hypopituitarism have an increased incidence of asymptomatic premature atherosclerosis and excess mortality because of cardiovascular disease. Although untreated growth hormone deficiency has been the most commonly implicated etiologic factor, the effect of other hormone replacement regimens should be considered. Short-term studies in adults attempting to relate hydrocortisone dose to a tissue effect have found increased bone turnover markers in hydrocortisone overreplaced patients and changes in carbohydrate metabolism on “conventional” (ie, 30 mg per day, higher than currently recommended) but not on lower-dose hydrocortisone regimens (20 mg per day).

Two studies have assessed hydrocortisone replacement in children with ACTH deficiency. Twice-daily (BD) hydrocortisone regimens resulted in supraphysiologic postdose cortisol peaks and low predose nadirs despite a relatively high mean daily hydrocortisone dose (12.3 mg/m² per day) in 44 patients with craniopharyngioma. In another study, although more satisfactory cortisol levels were described using limited filter paper sampling on BD therapy (mean dose: 8.9 mg/m² per day), overnight data and glucose levels were not collected.

The rationale for choosing thrice-daily (TDS) rather than BD regimens was because of the study by Groves et al, which showed that TDS regimens prevented afternoon hypocortisolemia and may improve well-being in adrenally insufficient adults. There are no pediatric studies in hypopituitarism of TDS hydrocortisone regimens despite its common usage, nor have cortisol profiles been reported with either BD or TDS regimens using lower hydrocortisone doses.

Children with hypopituitarism and their parents often report lethargy and difficulty concentrating, especially in the afternoon. To date, no studies have attempted to quantify the cognitive changes associated with acute fluctuations in plasma cortisol and symptomatology in children and adolescents with hypopituitarism. Clinical assessment of hydrocortisone replacement is generally based on assessment of well-being, questioning regarding symptoms of glucocorticoid deficiency, physical examination for signs of overreplacement, and dose adjustments based on current body surface area in the growing child. Because cortisol profiles are invasive and labor intensive, detailed assessments of glucocorticoid replacement regimens are not routinely performed in pediatric practice. Using 24-hour cortisol and glucose profiles along with computerized cognitive testing, our aim was to assess the currently prescribed hydrocortisone regimens in children and adolescents with hypopituitarism and to refine prescribing recommendations (including administration times and dose distribution). We hypothesized that TDS dosing was more physiological than BD dosing.

Subjects and Methods
Subjects
The study population consisted of 20 children and adolescents (aged 2.9–18.5 years) with hypopituitarism who were attending The Children’s Hospital at Westmead Endocrine Clinic. Any patient diagnosed previously with ACTH deficiency was eligible for the study. Patients who had required stress dose steroids in the previous 3 months were excluded. Other pituitary hormone deficiencies were satisfactorily replaced with thyroxine (n = 18), biosynthetic human growth hormone (GH; n = 11), intranasal desmopressin acetate (n = 7), oral desmopressin acetate (n = 4), oral estradiol valerate (2 girls), and oral testosterone undecanoate (3 boys). No patient had hyperprolactinemia. Five GH-deficient patients who were growing satisfactorily without GH therapy or had achieved near final height were not receiving GH therapy.

To assess the hydrocortisone regimens in the patients with hypopituitarism, plasma cortisol and glucose profiles were assessed over a 24-hour period. Clinical symptoms of hypocortisolemia and cognitive function prehydrocortisone and posthydrocortisone dose were also assessed.

Data from 22 healthy siblings aged 5.1 to 18.5 years were used to define normal 24-hour plasma cortisol profiles. The characteristics of the study subjects and control subjects are detailed in Table 1.

The study was approved by the Ethics Committee of The Children’s Hospital at Westmead. All of the participants and their parents gave informed consent/assent after receiving a detailed verbal explanation and written information about the study.

Experimental Design
Patients were admitted at 7:00 AM to the Endocrine Testing Unit at The Children’s Hospital at Westmead. Local anesthetic cream was applied to the cannulation site at 7:00 AM, and intravenous cannulation was per-
formed at 7:30 AM. Venous blood samples for plasma cortisol and glucose were collected into lithium heparin tubes via the previously placed intravenous cannula to minimize any distress from venipuncture. Blood samples were collected immediately before each dose of hydrocortisone to capture the plasma cortisol nadir, 1 hour after hydrocortisone to capture the plasma cortisol peak, and at other defined time points throughout the 24-hour period. Samples were refrigerated immediately and plasma separated by centrifugation within 6 hours. Plasma was stored at −80°C until batch analysis.

The hydrocortisone-administration and blood-sampling times are detailed in Table 2. Subjects fasted from supper (9:00 PM) until just after their 8:00 AM blood sample the next morning. No other restrictions were placed on the patients or control subjects between sampling times. They were encouraged to follow their normal diet throughout the day.

Anthropometric measurements were obtained when subjects were without shoes and in light clothing. Height was measured by using a wall-mounted Harpenden stadiometer (Holtain Ltd, Crymych, United Kingdom) and was accurate to 0.1 cm. Weight was measured by electronic scales and was accurate to 0.1 kg.

Laboratory Assays
Total cortisol concentration in venous plasma was evaluated using an Immulite 1000 cortisol chemiluminescence immunoassay (Diagnostic Products Corp, Los Angeles, CA). The mean interassay and intraassay coefficients of variation for the cortisol assay were 6.0% and 4.1%, respectively. The lower detection limit for the cortisol assay was 10 nmol/L (coefficient of variation: 11%; plasma cortisol conversion factor: 1/27.625 nmol/L). Reference ranges for plasma cortisol levels were derived by using the 2.5th and 97.5th percentiles of control-subject cortisol values at each time point. Glucose was measured on a VITROS Fusion using dry-slide chemistry (Ortho-Clinical Diagnostics, Raritan, NJ).

Other Data Collected
Patient- and parent/carer-reported symptoms of lethargy, weakness, concentration difficulties, or other symptoms at any time point throughout the day were recorded. Patients were asked which of these symptoms (if any) they attributed to their hydrocortisone dose, timing, or cortisol levels.

In ACTH-deficient patients ≥8 years old, cognitive function was assessed immediately before and 1 to 2

### TABLE 1  ACTH-Deficient Study Patient and Control-Subject Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>ACTH-Deficient Patients</th>
<th>Control Subjects (n = 22)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hydrocortisone BD (n = 11)</td>
<td>Hydrocortisone TDS (n = 9)</td>
</tr>
<tr>
<td>Age, median (range), y</td>
<td>12.7 (5.1 to 18.5)</td>
<td>7.9 (2.9 to 15.8)</td>
</tr>
<tr>
<td>Female, n</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Etiology, n</td>
<td>Congenital</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Acquired</td>
<td>7</td>
</tr>
<tr>
<td>Height SDS, median (range)</td>
<td>−0.4 (−1.5 to 0.5)</td>
<td>−0.8 (−4.7 to 1.4)</td>
</tr>
<tr>
<td>Weight SDS, median (range)</td>
<td>0.1 (−0.1 to 1.8)</td>
<td>0.1 (−2.9 to 3.4)</td>
</tr>
<tr>
<td>BMI SDS, median (range)</td>
<td>0.5 (−0.4 to 1.8)</td>
<td>1.3 (−0.1 to 2.6)</td>
</tr>
<tr>
<td>Hydrocortisone dose, median (range), TDD per m²</td>
<td>7.3 (4.4 to 11.4)</td>
<td>7.3 (4.7 to 12.8)</td>
</tr>
<tr>
<td>Pituitary deficiencies, n</td>
<td>GH</td>
<td>11</td>
</tr>
<tr>
<td>Thyrotropin</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>LH/FSH</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>ADH</td>
<td>5</td>
<td>7</td>
</tr>
</tbody>
</table>

GH indicates luteinizing hormone; FSH, follicle-stimulating hormone; ADH, antidiuretic hormone; NA, not applicable; SDS, SD score.

### TABLE 2  Hydrocortisone Regimens and Plasma Cortisol- and Glucose-Sampling Times

<table>
<thead>
<tr>
<th>Dosing Time</th>
<th>Sampling Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00 AM</td>
<td>12:00 PM</td>
</tr>
<tr>
<td>TDS</td>
<td></td>
</tr>
<tr>
<td>8:00 AM, 2:00 PM, 8:00 PM</td>
<td>X</td>
</tr>
<tr>
<td>8:00 AM, 4:00 PM, 8:00 PM</td>
<td>X</td>
</tr>
<tr>
<td>BD</td>
<td></td>
</tr>
<tr>
<td>8:00 AM, 4:00 PM</td>
<td>X</td>
</tr>
<tr>
<td>8:00 AM, 8:00 PM</td>
<td>X</td>
</tr>
<tr>
<td>Control subjects</td>
<td>X</td>
</tr>
</tbody>
</table>

X indicates when plasma cortisol and glucose samples were collected.

* Data are for the cortisol sample only.
hours after each dose of hydrocortisone. A computerized
cognitive test battery from CogState (CogState Ltd, Carlton
South, Victoria, Australia) was administered in a
quiet room. The test battery consisted of 4 tasks in the
form of card games on a laptop computer and assessed
psychomotor reaction time, memory (1 card learning
task), and executive function (strategy learning task and
choice/decision-making task). Each patient had 1 prac-
tice attempt (first administration) before commencing
the study test (second administration). The tests were
chosen because they are quick and easy to administer
(12 minutes to complete test battery), have no signifi-
cant practice effect after the second administration, have
been validated in children as young as 8 years old,16 and
have sensitivity to detect impairment of motor ability,
memory, attention, and concentration. For each task,
psychomotor reaction time was measured in millisec-
donds. In addition, the number of correct responses was
recorded and expressed as a percentage of the total
number of trials for each task.

Statistical Analysis
An independent sample t test was used to compare base-
line characteristics and total daily dose (TDD) of hydro-
cortisone between different treatment groups. Charac-
teristics of the cortisol profiles on BD and TDS hydro-
cortisone regimens were compared (hours below or
above the reference range). For this analysis, the
following definitions were used: waking hours = 8:00
AM to 8:00 PM, sleeping hours = 8:00 PM to 8:00 AM,
prolonged hypocortisolemia = plasma cortisol level of
<50 nmol/L for ≥4 hours, complete ACTH deficiency =
8:00 AM cortisol level of <100 nmol/L, partial ACTH
deficiency = 8:00 AM cortisol level of ≥100 nmol/L, and
hypoglycemia = blood glucose level of <3.5 mmol/L.
For normally distributed data, a Mann-Whitney U test
was used to compare the treatment regimens. One sam-
ple Student’s t test was used to compare predose and
postdose cognitive function. A nonparametric, 1-sample
Kolmogorov-Smirnov test was used for analysis of non-
normally distributed data. Spearman’s rank-order corre-
lation coefficient was used to assess the relationship
between cortisol dose, cortisol peak, time to nadir, and
cortisol area under the curve (AUC).

RESULTS
Twenty ACTH-deficient patients participated (11 girls, 7
with congenital hypopituitarism). The oral hydrocortisone
dosing regimen was BD in 11 and TDS in 9 patients. The
mean TDD of hydrocortisone on BD regimens was 7.6 ±
2.1 mg/m² per day and on TDS regimens was 8.3 ± 2.6
mg/m² per day. See Table 1 for baseline characteristics
of the control group and patient group by treatment regimen.

Comparison of Regimens
Hydrocortisone replacement regimens resulted in corti-
sol profiles outside the control range in all 20 of the
patients taking oral hydrocortisone (see Fig 1). Compar-
ing the BD and TDS regimens, the TDD of hydrocorti-
sone, total number of hours above the reference range,
and sleeping hours (8:00 PM to 8:00 AM) below the ref-
ERENCE range were similar (see Table 3). However, those
on the BD dose had more waking hours (8:00 AM to
8:00 PM) below the reference range than those on the
TDS dose (5.5 vs 2.1; P = .006). Nocturnal prolonged
hypocortisolemia (plasma cortisol level of <50 nmol/L
for ≥4 hours) was common on both regimens. Daytime
prolonged hypocortisolemia was more common on BD
than TDS regimens (see Table 3).

Cortisol AUC was >97.5th percentile for control sub-
jects in 3 patients and <2.5th percentile in 7 patients.
There was no significant difference in mean AUC for
those on BD and those on TDS hydrocortisone (3050.7
vs 3716.2 nmol/hour per L; P = .40; see Table 3). Ex-
cluding 4 patients (1 on TDS dose and 3 on BD dose)
with partial ACTH deficiency, hydrocortisone TDD per
square meter correlated with AUC (r = 0.77; P < .005).

All of the patients receiving oral hydrocortisone had
1-hour postdose plasma cortisol peaks >97.5th percen-
tile at some point during the 24-hour sampling period.
There was a modest correlation between hydrocortisone
dose and plasma cortisol peak, with wide variability in
peak cortisol levels for similar doses (see Fig 2A). Morn-
ing doses >4 mg/m² caused larger peaks than those <4
mg/m² (151 vs 47 nmol/L above the 97.5th percentile; 
P = .022). However, there was no difference in the length
of time taken to reach nadir below the 2.5th percentile
(5.2 vs 4.8 hours; P = .5). There was no significant
correlation between morning dose and time to reach
nadir (see Fig 2C). Similarly, night time doses >2.5
mg/m² resulted in larger cortisol peaks than doses <2.5
mg/m² (365 vs 148 nmol/L above the 97.5th percentile;
P = .05) but did not prevent early morning cortisol
nadir when compared with doses <2.5 mg/m² (7.1 vs
6.6 hours to fall below the 2.5th percentile; P = .19; see
Fig 2 B and D). Cortisol levels remained above the 2.5th
percentile for longer after the evening dose than after
the higher morning doses (2.8 ± 1.1 mg/m² vs 4.0 ± 1.2
mg/m², P = .001 and 6.9 ± 0.7 hours vs 5.3 ± 1.3 hours;
P < .005; see Fig 2 C and D).

Glucose Levels
No hypoglycemia (plasma glucose ≤ 3.5 mmol/L) was
detected at any sampling time point. One patient was
diagnosed with type 2 diabetes on the basis of the blood
glucose profile. She was receiving a relatively low dose
of hydrocortisone (TDD 5.7 mg/m² per day) and had a
family history of type 2 diabetes. Excluding this patient,
the mean fasting blood glucose levels at 4:00 AM (5.0 ±
0.5 mmol/L) and 8:00 AM (4.7 ± 0.4 mmol/L) did not
differ between those with plasma cortisol levels of <20
or >20 or 50 nmol/L cut points (see Table 4). There were no episodes of hyperglycemia (plasma glucose level of \( \geq 11.1 \) mmol/L) in association with 1-hour postdose cortisol peaks (mean blood glucose level: \( 5.2 \pm 0.8 \) mmol/L at 9:00 AM, \( 5.2 \pm 0.3 \) mmol/L at 3:00 PM, \( 4.9 \pm 0.4 \) mmol/L at 5:00 PM, and \( 5.4 \pm 0.5 \) mmol/L at 9:00 PM; see Table 5).

**Symptoms and Cognitive Function**

Ten patients described lethargy, which they attributed to low cortisol levels prehydrocortisone dose (BD dose: 5 patients; TDS dose: 5 patients). In the morning, the mean cortisol levels in symptomatic and asymptomatic patients were 16.7 and 9.9 nmol/L (\( P = .34 \)), when glucose levels were 4.5 and 5.1 mmol/L (\( P = .94 \)). In the evening, cortisol levels were 62.6 nmol/L in symptomatic vs 63.8 nmol/L (\( P = .98 \)) in asymptomatic patients when blood glucose levels were 5.1 and 5.3 mmol/L (\( P = .54 \)).

On predose and postdose cognitive testing (34 paired tests), no significant change in reaction speed was detected in the whole group (453.3 vs 438.8 milliseconds; \( P = .44 \)) or in subgroup analysis of those who had symptoms of predose lethargy, predose cortisol level of

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hydrocortisone Regimen</th>
<th>BD (n = 11)</th>
<th>TDS (n = 9)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hours above reference range, mean</td>
<td></td>
<td>6.3</td>
<td>7.4</td>
<td>.43</td>
</tr>
<tr>
<td>Waking</td>
<td></td>
<td>2.2</td>
<td>3.6</td>
<td>.27</td>
</tr>
<tr>
<td>Sleeping</td>
<td></td>
<td>4.2</td>
<td>3.9</td>
<td>.59</td>
</tr>
<tr>
<td>Hours below reference range, mean</td>
<td></td>
<td>9.8</td>
<td>7.1</td>
<td>.13</td>
</tr>
<tr>
<td>Waking</td>
<td></td>
<td>5.5</td>
<td>2.1</td>
<td>.006</td>
</tr>
<tr>
<td>Sleeping</td>
<td></td>
<td>4.4</td>
<td>4.6</td>
<td>.87</td>
</tr>
<tr>
<td>Patients with PH, n (%)</td>
<td></td>
<td>9 (82)</td>
<td>7 (78)</td>
<td>.82</td>
</tr>
<tr>
<td>Waking</td>
<td></td>
<td>7 (65)</td>
<td>0 (0)</td>
<td>.003</td>
</tr>
<tr>
<td>Sleeping</td>
<td></td>
<td>8 (73)</td>
<td>7 (78)</td>
<td>.79</td>
</tr>
<tr>
<td>Cortisol AUC, mean, nmol/hour per L*</td>
<td></td>
<td>3050.7</td>
<td>3716.2</td>
<td>.40</td>
</tr>
</tbody>
</table>

Waking indicates 8:00 AM to 8:00 PM; sleeping, 8:00 PM to 8:00 AM; Ph, prolonged hypocortisol-
<50 nmol/L, or prolonged predose hypocortisolemia. Predose and postdose tests of attention, learning, and memory also did not differ significantly.

**DISCUSSION**

This is the first study to report cortisol and blood glucose profiles on both BD and TDS regimens in children with hypopituitarism on the currently recommended lower daily dose of hydrocortisone. Our data confirm findings from adult studies that TDS dosing results in more phys-
iological cortisol profiles than BD. Despite this, all of the regimens produced unphysiological profiles reflecting the rapid absorption and short circulating half-life of hydrocortisone. The mean total daily dose in our group was less than that used in the only other large pediatric study (7.9 vs 12.3 mg/m² per day). This lower total daily dose was not associated with hypoglycemia despite the 11-hour overnight fast and the low cortisol levels for prolonged periods of time suggesting that these lower doses are unlikely to cause fasting hypoglycemia in children with hypopituitarism. A limitation of our study is that the youngest participant was 2.9 years old. Therefore, we did not study any infants or young toddlers with hypopituitarism who would be at higher risk of hypoglycemia and in whom the developing brain is more vulnerable to the effects of hypoglycemia.

To assess the study patients under relatively normal circumstances, they were asked to follow their usual dietary habits. It is possible that food intake could have masked hypoglycemia on the study day. Therefore, our study does not provide information about unusual or extreme events, such as a missed meal, a vomited meal, an intercurrent illness, or a change in routine.

From Fig 1, 4 patients with partial ACTH deficiency can be identified by their endogenous early morning cortisol rise. The results remained unchanged when these 4 patients were excluded from each analysis. However, it is possible that the small numbers of participants in this study may have prevented the detection of small differences between groups. For example, a retrospective power analysis revealed that 48 patients, 24 in each group (cortisol <50 or >50 nmol/L), would have been required to detect a lasting blood glucose difference of 0.5 mmol/L with 80% power.

Another study weakness may be that patients were studied on their prescribed dosing regimen rather than in a randomized crossover design. Regimens were chosen by the treating endocrinologist based on their personal prescribing habit. It is possible that individual patient variations in cortisol metabolism may have led to regimen changes before study entry. To address this possible bias, we reviewed each patient’s medical charts for information about previous hydrocortisone regimens. Three patients currently on BD hydrocortisone had previously been on TDS doses. These 3 changed regimen (from TDS to BD dose) because of the difficulties with administering a dose during school hours. Only 1 patient was changed from BD dose to TDS hydrocortisone on the basis of afternoon fatigue and reported an improvement in symptoms. Unfortunately, it is not possible to derive from this 1 case whether a placebo effect or a true biological effect was observed. The other study patients who complained of lethargy had not had dose adjustments made in response to their symptoms. Therefore, it is unlikely that previous regimen changes caused a significant bias in this study.

The fact that many of our patients with ACTH deficiency taking oral hydrocortisone had prolonged episodes of hypocortisolemia while maintaining seemingly normal activity, normal blood glucose levels, and denying symptomatology indicates that the intracellular half-life (ie, the intracellular processes initiated by hydrocortisone) continue for much longer than circulating plasma cortisol half-life would suggest. Although this is reassuring, important clinical questions remain unanswered. For example, do more physiologic profiles translate into improvements in functioning and cognition, when is dose adjustment required, and is there a better measure of cortisolemia than total plasma cortisol (eg, free plasma cortisol or direct measurement of in vivo glucocorticoid activity)?

Historically, dose adjustments in pediatric practice have been made on the basis of symptomatology and body surface area–related dose calculations. However, because symptomatology is often vague and may not be perceived by the patient, a more objective measure of symptomatic hypocortisolemia would be beneficial. In addition, in our patients with complete ACTH deficiency, predose cortisol levels did not differ between those with symptoms and those without symptoms. Therefore, we attempted to quantify symptomatic hypocortisolemia by assessing cognitive function prehydrocortisone and post-hydrocortisone dose. We were unable to find any difference in reaction speed, memory, concentration, or learning in our children with hypopituitarism. It is possible that the cognitive tests that we administered did not measure the appropriate cognitive variable, were not sensitive enough to detect the cognitive variability associated with fluctuations in cortisol, and/or subjective symptoms are unreliable indicators of symptomatic hypocortisolemia. These findings highlight the need for larger randomized studies addressing clinical and biological end points in children taking oral hydrocortisone.

Although all of the patients who were taking oral hydrocortisone experienced supraphysiological cortisol levels, from our data we can infer that increasing the dose per square meter above certain “threshold” levels (4 mg/m² in the morning and 2.5 mg/m² in the afternoon and evening) only worsens the magnitude of cortisol peaks without extending the length of time taken to fall below the reference range. Presumably this reflects rapid glomerular cortisol filtration and renal excretion, which occurs after each dose, when plasma cortisol levels are high and plasma cortisol binding sites are fully saturated.

Excluding the patient with type 2 diabetes, we did not detect any hyperglycemia during the supraphysiological cortisol peaks or at any time point. Our results are in agreement with data from McConnell et al. They reported no increase in hepatic or peripheral insulin resistance on lower-dose replacement therapy in adults.

The observation that, despite lower doses in the
evening, cortisol levels were maintained above the 2.5th percentile for longer than after the morning dose (6.9 ± 0.7 vs 5.3 ± 1.3 hours) can be explained by decreased cortisol clearance in the evening and overnight, which has been documented previously in other adult studies.20,18,19 In addition, there may have been differences in hydrocortisone absorption related to dose timing and food intake, which we did not control for.

There are more obstacles to accurate prescribing of hydrocortisone than the unanswered medical questions. As with other chronic illnesses in children and adolescents, administration of medication during the day can be logistically difficult. The currently available hydrocortisone preparations (4-, 5-, 10-, and 20-mg tablets) are not ideal. Although oral hydrocortisone suspension could make small dose adjustments easier, it was withdrawn because of concerns about poor bioavailability.20 Although longer-acting synthetic glucocorticoids (such as prednisolone, methylprednisolone, and dexamethasone) could prevent nadirs, blood levels are difficult to monitor, and their more potent glucocorticoid activity may make overreplacement likely. Even if a slow release hydrocortisone preparation were available, as is currently being developed in the United Kingdom,21 it is unlikely that the body surface area adjustments necessary in pediatric prescribing would be possible.

On the basis of these data, we would recommend aiming for a total daily hydrocortisone dose of 6 to 8 mg/m² per day, which equates to estimates of daily cortisol production and does not seem to be associated with altered glucose metabolism or cognitive changes. Caution should be taken in management of younger children (<3 years old) with ACTH deficiency who are at a higher risk of hypoglycemia. The regimen and dose splitting should be tailored to suit each individual child. Cortisol and glucose profiles may be useful in difficult-to-manage patients or those at higher risk of hypoglycemia. Additional larger studies to identify and quantify the symptomatology and cognitive defects associated with acute hypocortisolemia are required. Once identified, these may be useful tools to guide hydrocortisone dose adjustment.

ACKNOWLEDGMENTS

This study was supported by the following research grants (to Dr Maguire): National Health and Medical Research Council, Australian Government, Medical Postgraduate Scholarship; Pfizer Pediatric Endocrine Career Development Grant; and Eli Lilly Australasian Pediatric Endocrine Group Research Grant.

REFERENCES

Mass Distribution of Free, Intranasally Administered Influenza Vaccine in a Public School System

L. Rand Carpenter, DVMa,b, John Lott, MS5, Brian M. Lawson, MA5, Stephanie Hall, MD, MPH, Allen S. Craig, MD5, William Schaffner, MD5, Timothy F. Jones, MD

aEpidemic Intelligence Service, Centers for Disease Control and Prevention, Atlanta, Georgia; bCommunicable and Environmental Disease Services, Tennessee Department of Health, Nashville, Tennessee; cKnox County Health Department, Knoxville, Tennessee; dDepartment of Preventive Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee

Financial Disclosure: Mr Lott and Dr Jones were unpaid attendees at a meeting of the MedImmune 2006 FluMist for Schools Advisory Board meeting, where results of this mass vaccination campaign were presented. Dr Schaffner served as a session moderator at the same meeting and received an honorarium.

ABSTRACT

OBJECTIVE. School-based influenza vaccination programs are a potentially important method of protecting the community against influenza. We evaluated the feasibility and success of a large, school-based influenza vaccination campaign.

METHODS. On-site administration of intranasally administered, live attenuated influenza vaccine was offered to all students and staff members in a large, metropolitan public school system in October to December 2005. We evaluated vaccine coverage levels, resources expended, and physician and parent attitudes and knowledge.

RESULTS. Of 53,420 public school students, 24,198 were vaccinated with live attenuated influenza vaccine. Of 5,841 school staff members, 3,626 were vaccinated with live attenuated influenza vaccine or inactivated influenza vaccine. The proportions of students vaccinated were 56% among elementary schools, 45% among middle schools, and 30% among high schools. Schools with larger proportions of black or low-income families had lower vaccine coverage levels. The health department and school system expended 6,900 person-hours during the campaign, and various health department clinics were closed for a total of 84 half-days. Community physicians were supportive of the campaign and frequently advised participation for eligible patients. Some physicians had misunderstandings about live attenuated influenza vaccine contraindications. Concern about adverse effects, having asthma, negative physician advice, and nonparticipation in any vaccination program were common reasons for students not participating.

CONCLUSIONS. This influenza vaccination campaign in a large public school system achieved relatively high vaccine coverage levels but required a substantial resource commitment from the local health department. This evaluation has critical implications for the ongoing debate regarding immunization policies for school-aged children and preparedness plans for pandemic influenza.
Five percent to 20% of the US population may be infected with influenza each year, and influenza is associated with an average of ≥200,000 hospitalizations and 36,000 deaths annually.1–2 Among school-aged children, influenza attack rates can be as high as 25% to 30%, generally higher than for all other age groups.3–5 This high influenza morbidity rate among schoolchildren leads to excess school absenteeism and increased parental work absenteeism.6 Children have long been considered to be important sources of influenza in the community.4,6,7

Vaccination of schoolchildren in a small Michigan city reduced overall influenza morbidity in the community.8 A school vaccination program in Japan resulted in reduced numbers of influenza-related deaths among elderly individuals.9 Recent studies demonstrated reductions in the occurrence of acute respiratory illness in families and the community after vaccination of 25% to 47% of school-aged children with intranasally administered, live attenuated influenza vaccine (LAIV).10–12 LAIV is a temperature-sensitive, live attenuated virus vaccine that offers protection against influenza A and B viruses. This cold-adapted, heat-labile vaccine replicates readily in the upper respiratory tract and produces an immune response but does not replicate efficiently in the lower respiratory tract. The risk of LAIV virus transmission to others is thought to be very low.13–14 LAIV is licensed for use among healthy children and adults 5 to 49 years of age. LAIV has been reported to provide effective durable immunity, including protection against antigenically related influenza strains not included in the vaccine.15–17 It is administered annually as a single-dose nasal spray into each nostril and is well tolerated by children.18 These characteristics make LAIV an attractive choice for use in school-based mass vaccination campaigns. Strategies for mass vaccination have implications for planning for pandemic influenza. We evaluated a campaign to vaccinate students with LAIV in a large, metropolitan, public school system, to assess feasibility and success.

METHODS
Knox County, Tennessee (population: 389,327), includes the city of Knoxville and has a single public school system. The county health department staff of 299 persons includes 84 doctors and nurses. In June to December 2005, a campaign was conducted to provide LAIV to students and staff members in the Knox County public school system. The campaign’s objective was to offer LAIV, free of charge, to all eligible students ≥5 years of age in kindergarten through 12th grade and to school staff members. The vaccine was donated by the manufacturer (MedImmune, Gaithersburg, MD), but most costs associated with education, administration of the campaign, and purchase of inactivated vaccine were borne by the Knox County Health Department. Planning and community education began in June, and vaccinations occurred in October to December. Initial planning included managers from the health department and the school system. County health department staff members made numerous informational presentations to the regional children’s hospital staff, a local medical organization, school principals, and parent-teacher associations. News media were informed through press releases and interviews with health department staff members. Information about childhood influenza vaccination and LAIV was sent directly to local physicians’ offices and hospitals through the use of facsimile and mail. The health department established a telephone help-line for questions from parents, school staff members, students, and health care providers in the weeks before and during the campaign.

Health department staff members were responsible for LAIV inventory, including ordering, storing, and dispensing. Schools sent an introductory letter, a vaccine information sheet, the Centers for Disease Control and Prevention LAIV Information Statement, and a consent form home with each student. The completed forms, signed by a parent or guardian, were collected by school personnel; distribution and collection methods varied among schools. Vaccination teams, including health department and school nurses, health department physicians, and administrative personnel, were trained in campaign protocols. The first round of vaccination included 3 to 6 schools per day, with a target time limit at each location of 3 hours.

Recipients were required to be 5 to 49 years of age with no chronic medical conditions, immunosuppression, hypersensitivity to egg products, history of Guillain-Barré syndrome, or history of asthma, as reported on the consent form. School staff members who were not eligible for LAIV were offered trivalent inactivated influenza vaccine (TIV), but students were not offered TIV because of financial and time constraints during the campaign. Students <9 years of age who had not been administered influenza vaccine (either TIV or LAIV) previously were offered a second dose of LAIV. Second doses were provided in December, during a single follow-up visit to each school; at the same time, persons who desired vaccination but had not been vaccinated during the first round were offered LAIV.

Information regarding each school’s vaccine coverage levels and student and staff member demographic features was collected by health department staff members. Health department administrators tracked resources expended by health department and school nursing staff members throughout the campaign. Coverage levels were based only on vaccine administered on-site, and the number of persons vaccinated outside the campaign was not assessed. Eligibility for the National School Lunch Program, based on low household income, was used to approximate the mean socioeconomic status of

PEDIATRICS Volume 120, Number 1, July 2007 e173
each school’s students. School absenteeism after vaccination was monitored daily by health department personnel, using a previously described, automated, early aberration reporting system.18

An anonymous questionnaire was mailed to all primary care physicians in Knox County, to assess knowledge, attitudes, and perceptions regarding the campaign and influenza vaccination. A second survey asking about campaign participation of 1 student in the household was mailed to the parents or guardians of a random sample of students. The sample of 11% of students was selected by using the SPSS (SPSS Inc, Chicago, IL) “random sample of cases” function. Responses were returned anonymously. Data were analyzed by using Epi-Info (Centers for Disease Control and Prevention, Atlanta, GA) and SAS 9.0 (SAS Institute, Cary, NC) software. This evaluation was classified as a public health program evaluation by the institutional review board of the Tennessee Department of Health and the Centers for Disease Control and Prevention.

RESULTS

Vaccine Coverage

The school district consisted of 81 schools with total enrollment of 53 420 students, including 43 270 (81%) white, 8013 (15%) black, and 2137 (4%) other races/ethnicities, as classified by the school system. Fifty elementary schools (kindergarten to 5th grade), 14 middle schools (6th–8th grades), and 12 high schools (9th–12th grades) participated in the campaign. Five alternative schools for kindergarten to 12th-grade students with special needs, students with behavioral disorders, and adult-education students (total enrollment: 197 students) were not included in this analysis.

The campaign resulted in 24 198 (45%) students vaccinated with at least 1 dose of LAIV at school. Vaccination levels were highest among elementary students (56%) and lowest among high school students (30%) (Table 1). Vaccination levels in individual schools varied from 13% to 75%. Elementary school students represented 46% of all students in the school system and 57% of vaccinated students.

A second dose of LAIV was administered to 2945 (58%) of 5099 students <9 years of age who received LAIV during the primary vaccination day and who required a second dose. When vaccination levels were adjusted to reflect only students who were fully vaccinated, 47% of elementary students were fully vaccinated in this campaign.

The proportion of students at each school who were eligible for free or reduced-price lunches through the National School Lunch Program ranged from 6% to 99% (mean: 48%). Students in schools with relatively higher levels of enrollment in the National School Lunch Program were less likely to have been vaccinated than were students in schools with lower levels of enrollment in the lunch program (Fig 1). Similarly, students in schools with relatively higher levels of enrollment of black students were less likely to have been vaccinated than were students in schools with lower levels of enrollment of black students (Fig 2).

Approximately one half of the students did not return consent forms. In a convenience sample of 5 schools, a mean of 5.1% of each school’s returned forms indicated ineligible for LAIV.

Of 5841 school staff members, 3626 (62%) were vaccinated; 1464 (40%) received LAIV and 2162 (60%) were given TIV. The staff vaccination level in elementary schools was 68%, that in middle schools was 64%, and that in high schools was 58%.

No severe adverse reactions to LAIV were reported to the health department or school nurses. No marked change in absenteeism among students in any school was detected during the 2 weeks after vaccine administration.

Surveys

Questionnaires were mailed to 622 primary-care physicians. A total of 331 questionnaires (53%) were returned, and 268 respondents (81%) were aware of the vaccination campaign. Of the respondents, 233 (70%) practiced family medicine, pediatrics, or internal medicine. All pediatricians who responded were aware of the campaign. Of the 196 physicians who had given patients advice regarding the campaign, 185 (94%) had advised ≥1 patient to participate and 103 (53%) had advised ≥1 patient against vaccination. Reasons for advising a student against receiving LAIV included asthma (cited by 74% of respondents), immunocompromised status of the patient (34%), immunocompromised status of a close contact (34%), presence of a chronic metabolic disease (32%), and egg allergy or history of Guillain-

<table>
<thead>
<tr>
<th>School Type</th>
<th>Total No. of Students Enrolled</th>
<th>No. (%) Vaccinated</th>
<th>No. of Schools</th>
<th>Proportion of Students Vaccinated per School, Range, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elementary (kindergarten to grade 5)</td>
<td>24 769</td>
<td>13 809 (56)</td>
<td>50</td>
<td>26–75</td>
</tr>
<tr>
<td>Middle (grades 6–8)</td>
<td>12 500</td>
<td>5576 (45)</td>
<td>14</td>
<td>23–62</td>
</tr>
<tr>
<td>High (grades 9–12)</td>
<td>16 151</td>
<td>4813 (30)</td>
<td>12</td>
<td>13–51</td>
</tr>
<tr>
<td>Total</td>
<td>53 420</td>
<td>24 198 (45)</td>
<td>76</td>
<td>13–75</td>
</tr>
</tbody>
</table>
Barreé syndrome (27%). Eight physicians reported advising a patient (student or staff member) to be absent from the school on vaccination day, to avoid exposure to the vaccine virus. By using a scale of 1 to 5 (“not important” to “very important”), physicians ranked the importance of student influenza vaccination to the students’ health and to their families’ and community’s health; 84% ranked importance to the individual as either 4 or 5, and 92% ranked importance to the community as 4 or 5. All except 4 respondents ranked importance to the family and community the same as or higher than importance to the individual student.

A parent questionnaire was mailed to the homes of 5749 students. Responses were received from 1432 (25%), representing 2.7% of the student population. The distribution of respondents matched the proportions of the school population enrolled in elementary, middle, and high schools. Black students’ parents represented 5% of survey respondents, whereas black students represented 14% of the school system enrollment. Of survey respondents, 62% reported that their child had been vaccinated, compared with 45% vaccine coverage overall in the campaign. Nonparticipation in the vaccination campaign was reported by 34 (53%) of 64 parents of black students and 494 (36%) of 1339 parents of non-black students (risk ratio: 1.44; 95% confidence interval: 1.13–1.83). The most common reasons parents gave for students not being vaccinated in the campaign included
concerns about adverse effects (29%), having asthma (23%), being vaccinated elsewhere (19%), being advised by a physician to not participate (11%), and being someone who does not participate in vaccinations at all (10%).

Resources
The Knox County health officer and nursing director expended ~840 hours during 7 months (30% of full-time employment) preparing for and supervising the campaign. Health department staff members expended 4200 person-hours, and school nursing staff members contributed 2700 person-hours during the 2 months of vaccine administration, representing 85 person-hours per school. Approximately 75% of these person-hours were expended by professional medical personnel. Administrative hours contributed by school staff members in handling consent forms were not recorded.

To provide staffing for school vaccination days, certain health department services were closed periodically. Four remote-location health clinics and certain other health department services (eg, adult preventive care and indigent primary-care clinics) were closed for a total of 84 half-days during the 5-week vaccination campaign. Approximately 9900 patient visits were missed or delayed because of these closures.

Expenditure of health department funds for the campaign was approximately $43,000, of which $28,000 was used for the purchase of TIV. Freezers and alarm equipment for vaccine storage, dry ice for vaccine transport, printed information packets and consent forms, and expendable materials for vaccine administration cost approximately $15,000, a portion of which was supported by the LAIV manufacturer.

DISCUSSION
This school-based influenza vaccination campaign using LAIV achieved 45% coverage of students in kindergarten through 12th grade in a large, diverse, metropolitan school system. This evaluation demonstrates that coverage levels comparable to those of smaller pilot campaigns are achievable in large school systems with the application of substantial time and resources. Despite intensive efforts at education and promotion, certain population groups had low participation rates.

The LAIV package insert states that individuals with a history of asthma or reactive airway disease should not receive LAIV. The Advisory Committee on Immunization Practices also recommends that persons with asthma or reactive airway disease should not be vaccinated with LAIV. It is estimated that 5% to 15% of US children have a history of asthma and that self-reported and proxy-reported asthma rates may exceed 20%. By using a conservative 20% ineligibility estimate, the overall vaccine coverage level achieved among eligible students approached 60%. Previous LAIV vaccination efforts, with coverage levels as low as 20% of eligible children, demonstrated a reduction in the community’s burden of influenza-related illness. Stochastic modeling predicts that vaccination of only 20% of US children could reduce the number of influenza cases in the general population by 46%.

Influenza vaccination of children as a strategy for community control of influenza is a topic of substantial current interest. Several lines of evidence suggest that vaccinating younger and healthier groups, in addition to persons at high risk, may benefit the larger community as well as those being vaccinated. Japanese officials began a program of influenza vaccination in schools in 1962. In 1977, influenza vaccination of schoolchildren 7 to 15 years of age became mandatory; this was followed by remarkable nationwide declines in both total excess pneumonia and influenza mortality and all-cause mortality rates. After the vaccination requirements were repealed in 1994, excess pneumonia and influenza mortality rates increased again. Ecologic evaluation of the project suggests that the vaccination of schoolchildren had an effect on mortality rates in the older population.

In 1968 and 1969, during the Hong Kong influenza pandemic, a school-based influenza vaccination campaign (kindergarten through 12th grade) was conducted in the small community of Tecumseh, Michigan, and achieved an 86% vaccination level.

In a multisite study in 2004 and 2005, LAIV was administered to 2717 (47%) of 5840 students in 11 elementary or parochial schools. Additional schools at the same sites served as control schools. During peak influenza season, intervention school household members reported fewer episodes of influenza-like illness, purchased fewer medications, and missed fewer days of work and school than reported among control school household members. Absenteeism during the peak influenza period was reduced among vaccinated children, compared with nonvaccinated children, in the intervention schools. Similarly, a 3-year LAIV clinical trial in Texas demonstrated that vaccination of 20% to 25% of eligible children (age: 18 months to 18 years) in the community was associated with a modest reduction in the community burden of medically attended, acute respiratory illness.

A simultaneous study of the same LAIV-vaccinated children demonstrated protection against both influenza A and influenza B among vaccinees. Assessment of the feasibility of large-scale,
school-based interventions using LAIV is an important addition to these studies and was the focus of our evaluation.

Within each grade category, participation levels varied substantially among schools in Knox County. Socioeconomic status and race were important factors in this variability, and schools with relatively large proportions of economically disadvantaged students or black students had lower participation rates. The parent survey indicated that concern regarding vaccine adverse effects was a major barrier to participation. This finding is consistent with a 2004 study among Tennessee residents that indicated that nonwhite persons were less likely to be vaccinated for influenza because of concern regarding adverse effects and the lack of perceived need for influenza vaccination. Although the low response rate in our parent survey might have limited its representativeness, the survey results clearly point to a need to focus educational efforts among lower-participating groups to address safety concerns even more prominently. Fewer than 60% of children <9 years of age for whom a second dose was indicated received that second dose at school. It is possible that the second dose was delivered in another setting or that the importance of this second vaccine dose was not communicated adequately.

School district officials were eager to reduce absenteeism among students and staff members to meet state standards and to avoid school closures. During the previous year’s influenza season (2004–2005), the entire Knox County school system was closed for 2 days because of high staff absenteeism rates. To encourage support for this campaign, the health department purchased and provided TIV for staff members who were not eligible to receive LAIV. Because of financial and logistic concerns, TIV was not offered to students at school during the school-based campaign. Although it was available free to students at the Knox County Health Department, not offering TIV to LAIV-eligible students at school likely decreased vaccine coverage among certain groups of students, including children who reported a history of asthma.

Concern regarding vaccine virus transmission was a notable barrier to campaign participation. The safety of LAIV in different situations is well established, and multiple studies have detected minimal or no vaccine virus transmission. Knox County Health Department staff members made extensive efforts to educate physicians and parents regarding vaccine safety, benefits, and contraindications during the campaign, and the physician survey revealed an extremely positive response to the campaign among physicians. Nevertheless, some physicians demonstrated confusion regarding LAIV contraindications. Concerns about vaccine virus transmission prompted several physicians to advise some patients not to participate in the campaign. Reasons cited included having a household contact with asthma or a history of chronic disease, which are not valid contraindications for LAIV. At least 4 local obstetricians advised pregnant patients to be absent from school on the day of vaccination, to avoid exposure to vaccine virus.

This campaign required extraordinary resource commitment by both the health department and the school system. Motivated supervisors within the health department provided strong leadership. However, even with donated vaccine, the demands on health department and school system personnel and the effect on regular patient services made the vaccine campaign an expensive and disruptive endeavor.

Plans for future campaigns should include provisions to reduce the public health resource burden. Planners might consider using temporary staff members or limiting the campaign to younger students, to minimize the impact on regular health department and school health functions. In addition, it is important to improve targeted education for groups with low vaccination coverage and to focus on improved professional communication with local physicians, to correct misconceptions regarding LAIV.

CONCLUSIONS

We evaluated a widely accepted, resource-intensive intranasal influenza vaccine campaign in a large school system with 45% overall vaccine coverage among students. Certain groups, including high school students, black students, and students in schools with low family incomes, should be targeted for promotion and education during future campaigns, to improve acceptance of the vaccine. The substantial expenditure of resources necessary for this campaign might limit feasibility in other settings. However, vaccination of school-aged children could serve as an important strategy for reducing influenza morbidity in the wider community, and it will be important for future studies to estimate the effects of such campaigns on community influenza rates and influenza complications. This evaluation provides important information regarding the feasibility of school-based vaccination programs for children and has potential implications for pandemic preparedness plans for vaccine distribution.

ACKNOWLEDGMENTS

The Knox County Health Department received financial support and LAIV for this campaign from MedImmune. We thank Mike Winstead, PhD, Director of Research and Evaluation, Knox County Schools, for assistance with the parent survey.

REFERENCES

2. Thompson WW, Shay DK, Weintraub E, et al. Influenza-


Rapid-Onset Obesity With Hypothalamic Dysfunction, Hypoventilation, and Autonomic Dysregulation Presenting in Childhood

Diego Ize-Ludlow, MD⁷, Juliette A. Gray, BSc⁵, Mark A. Sperling, MD⁷, Elizabeth M. Berry-Kravis, MD, PhD⁶,⁷, Jeff M. Milunsky, MD⁶,⁷, I. Sadaf Farooqi, MD, PhD⁸, Casey M. Rand, BSc⁵, Debra E. Weese-Mayer, MD⁹

⁷Department of Pediatrics, University of Pittsburgh, Pittsburgh, Pennsylvania; ⁵University Department of Clinical Biochemistry, Addenbrooke’s Hospital, Cambridge, England; Departments of ⁶Pediatrics, ⁷Neurology, and ⁸Biochemistry, Rush Children’s Hospital, Rush University Medical Center, Chicago, Illinois; ⁹Center for Human Genetics and Departments of ⁶Pediatrics and ⁷Genetics and Genomics, Boston University School of Medicine, Boston, Massachusetts

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

ABSTRACT

OBJECTIVE. The goal was to characterize the phenotype and potential candidate genes responsible for the syndrome of late-onset central hypoventilation with hypothalamic dysfunction.

METHODS. Individuals with late-onset central hypoventilation with hypothalamic dysfunction who were referred to Rush University Medical Center for clinical or genetic assessment in the past 3 years were identified, and medical charts were reviewed to determine shared characteristics of the affected subjects. Blood was collected for genetic testing of candidate genes (PHOX2B, TRKB, and BDNF) and for high-resolution conventional G-bandning, subtelomeric fluorescent in situ hybridization, and comparative genomic hybridization analysis. A subset of these children were studied in the Pediatric Respiratory Physiology Laboratory at Rush University Medical Center.

RESULTS. Twenty-three children with what we are now naming rapid-onset obesity with hypothalamic dysfunction, hypoventilation, and autonomic dysregulation were identified. Comprehensive medical charts and blood for genetic testing were available for 15 children; respiratory physiology studies were performed at Rush University Medical Center on 9 children. The most characteristic manifestations were the presentation of rapid-onset obesity in the first 10 years of life (median age at onset: 3 years), followed by hypothalamic dysfunction and then onset of symptoms of autonomic dysregulation (median age at onset: 3.6 years) with later onset of alveolar hypoventilation (median age at onset: 6.2 years). Testing of candidate genes (PHOX2B, TRKB, and BDNF) revealed no mutations or rare variants. High-resolution chromosome analysis, comparative genomic hybridization, and subtelomeric fluorescent in situ hybridization results were negative for the 2 patients selected for those analyses.

CONCLUSIONS. We provide a comprehensive description of the clinical spectrum of rapid-onset obesity with hypothalamic dysfunction, hypoventilation, and auto-
nomic dysregulation in terms of timing and scope of symptoms, study of candidate genes, and screening for chromosomal deletions and duplications. Negative \textit{PHOX2B} sequencing results demonstrate that this entity is distinct from congenital central hypoventilation syndrome.

**Temperature Control, Energy Regulation, Cardiorespiratory Regulation, Thirst, and Water Balance**

Some of the most primitive functions of our nervous system are regulated by the basic circuitry involved in these processes. Although some of the neurological and neuroanatomical factors involved in conveying these regulatory signals are now known, knowledge of the basic circuitry involved is still limited. The identification of paired-like homeobox 2B (\textit{PHOX2B}) as the disease-causing gene in congenital central hypoventilation syndrome (CCHS) has enabled exploration of some of these ancestral functions as they relate to the autonomic nervous system. As CCHS gains visibility, other, seemingly overlapping diseases can be distinguished. Late-onset central hypoventilation syndrome (LO-CHS) with hypothalamic dysfunction (HD) is a less well known but potentially related condition of autonomic dysregulation.\textsuperscript{4} LO-CHS/HD was first described in 1965.\textsuperscript{5} LO-CHS/HD has a variable presentation including the following constellation of symptoms: hyperphagia and obesity, alveolar hypoventilation, altered respiratory control, thermal dysregulation, water imbalance, pain hyposensitivity, behavioral disorders, strabismus, pupillary anomalies, hyperprolactinemia, altered onset of puberty, and tumors of neural crest origin.\textsuperscript{5–14} In 2000, Katz et al\textsuperscript{6} reported on 1 child and provided a comprehensive review of 10 previously published cases of children with LO-CHS. One more case was reported subsequently.\textsuperscript{15} Katz et al\textsuperscript{6} suggested that LO-CHS/HD, an entity that presents exclusively after infancy, is distinct from CCHS. The classic cases of \textit{PHOX2B} mutation-confirmed CCHS present in the newborn period with alveolar hypoventilation in the absence of primary lung, cardiac, or neuromuscular abnormalities or an identifiable brainstem lesion that can account for the hypoventilation.\textsuperscript{20}

However, a growing number of individuals with \textit{PHOX2B} mutation-confirmed CCHS are being identified with presentation in childhood and adulthood,\textsuperscript{16,21–23} in contrast to the typical presentation in the newborn period. Similarly, the CCHS phenotype includes autonomic nervous system dysregulation,\textsuperscript{24} with tumors of neural crest origin\textsuperscript{20} and endocrine abnormalities, including growth hormone deficiency and hypothyroidism (unpublished data), in a subset of children.

There seems to be overlap in the clinical presentation of LO-CHS/HD and CCHS, but introduction of \textit{PHOX2B} as the disease-defining gene for CCHS allows for genetic distinction of LO-CHS/HD and CCHS and the opportunity to distinguish the 2 disease entities more clearly.

Two other candidate genes that might account for the phenotype of LO-CHS/HD encode the neurotrophin brain-derived neurotrophic factor (BDNF) and its receptor, TRKB, both of which are involved in neuronal development and maintenance and synaptic plasticity. Furthermore, evaluation of numerical and structural chromosome rearrangements with high-resolution, metaphase, comparative genomic hybridization (CGH) for patients with neurocognitive delays, with or without congenital anomalies/unusual phenotypes, has proved useful for uncovering the etiology and demonstrating interstitial chromosome deletions and duplications as small as 3 megabases.\textsuperscript{25}

Taken together, the phenotype of LO-CHS/HD suggests a cohesive entity that might have a genetic basis. In an effort to describe more fully the phenotype of LO-CHS/HD, we sought (1) to identify all children with symptoms consistent with this diagnosis who were referred for physiologic study or genetic testing in order to provide clear diagnostic criteria, (2) to rename the disease, to improve and to expedite patient identification and treatment, (3) to provide recommendations for testing, to confirm the diagnosis and to optimize long-term follow-up monitoring, and (4) to conduct candidate gene analysis.

**METHODS**

**Case Identification**

Individuals with clinical features consistent with LO-CHS/HD who were referred to Rush University Medical Center (RUMC) for clinical or genetic assessment between 2002 and 2006 were identified for inclusion in this institutional review board-approved study. Medical charts for each proband were reviewed by 2 authors (Drs Ize-Ludlow and Weese-Mayer), to determine the shared characteristics of LO-CHS/HD. Additional testing or records were requested on a case-by-case basis, to ensure uniform clinical assessments. Some of the children were studied in the Pediatric Respiratory Physiology Laboratory at RUMC.

**Criteria for Preliminary Diagnosis**

Our criteria for LO-CHS/HD included onset of alveolar hypoventilation after the age of 2 years and evidence of HD, as defined by ≥1 of the following findings: rapid-onset obesity, hyperprolactinemia, central hypothyroidism, disordered water balance, failed growth hormone stimulation test, corticotropin deficiency, or delayed or precocious puberty.

**Medical Chart Review**

All available clinical events, physical examination results, ancillary studies, laboratory data, diagnoses, and treatments were added to a deidentified database. The database was examined for patterns in presentation.
Height, weight, and BMI percentiles were calculated by using Epi Info 3.3.2 (Centers for Disease Control and Prevention, Atlanta, GA).

DNA Analyses

Sample Preparation and Testing Procedures
DNA from blood samples was prepared according to standard procedures and screened for the polyalanine expansion mutation in PHOX2B that is characteristic of CCHS. If this standard clinical test revealed the normal genotype coding for 20 alanine repeats on each allele, then institutional review board-approved, informed consent was obtained for sequencing of the entire coding sequence and intron-exon boundaries of PHOX2B, as well as other candidate genes. Genotyping of the PHOX2B polyalanine repeat, sequencing of the PHOX2B gene, and screening for mutations in BDNF and NTRK2 (the gene encoding TRKB) were performed as described by Weese-Mayer et al., Garcia-Barcelo et al., and Gray et al., respectively. In a subset of cases, high-resolution chromosome analysis, high-resolution metaphase CGH, and subtelomeric fluorescent in situ hybridization (FISH) analysis were performed.

High-Resolution Chromosome Analysis
Harvesting and G-banding were performed according to standard procedures. Twenty metaphases were analyzed at a resolution exceeding 600 bands.

High-Resolution Metaphase CGH
Slides with normal lymphocyte metaphase chromosomes for CGH analysis were stored at −20°C before hybridization. CGH was performed as described by Kirchhoff et al. The CGH hybridization slides were analyzed by using CytoVisionSystem 2.72 high-resolution CGH analysis software (Applied Imaging, Santa Clara, CA). Ten to 15 metaphases were captured by using a Zeiss fluorescence microscope with an integrating charge-coupled device camera (Photometrics, Tucson, AZ). The green (patient DNA) to red (reference DNA) fluorescence ratio along the length of the chromosomes was calculated. The standard reference interval was based on an average of normal cases, as described by Kirchhoff et al. The intervals were scaled automatically to fit the test case. The mean ratio profile of each case, with 99.5% confidence intervals, was compared with the average ratio profile of the normal cases, with similar confidence intervals.

Subtelomeric FISH Analysis
FISH with a subtelomeric DNA probe panel specific for the subtelomeric ends of each chromosome arm was performed according to the manufacturer’s instructions (Abbott Molecular, Des Plaines, IL). Computer-assisted analyses were performed by using a Zeiss Axioskop 2 fluorescence microscope with an integrating charge-coupled device camera (Photometrics).

RESULTS

Case Identification
We identified 23 referred children with clinical features consistent with LO-CHS/HD. Comprehensive medical charts were available for 15 of those children (6 male subjects and 9 female subjects), of whom 9 were studied in the Pediatric Respiratory Physiology Laboratory at RUMC. Referrals were primarily from pediatric pulmonologists, because of hypoventilation without recognition of the associated abnormalities.

Phenotype Data for LO-CHS/HD

Prevalent Manifestations
Review of the 15 cases with comprehensive medical charts revealed many characteristic features that occurred in all patients, as well as rare features that occurred in only a few patients. The age at onset and the frequency of each symptom are provided in Tables 1 and 2, respectively. The most prevalent manifestations were overweight (BMI of >95th percentile) with rapid-onset obesity and alveolar hypoventilation, followed in frequency by ophthalmologic manifestations, gastrointestinal dysmotility, and thermal dysregulation. The temporal relationships of these phenotypic features to one another are presented in Figs 1 and 2. The earliest manifestations were hypothalamic (median age at onset: 3 years), followed by autonomic (median age at onset: 3.58 years), behavioral (median age at onset: 4.8 years), and then respiratory (median age at onset: 6.17 years). All data in the sections that follow pertain to the 15 patients for whom comprehensive data were available. In the respiratory section, however, findings for the 15 patients are reviewed and then physiologic studies for the 9 children who were evaluated comprehensively at RUMC are reported.

HD
Evidence of HD was found for all 15 patients for whom complete medical charts were available. For 12 patients (80%), the initial symptom was rapid-onset obesity, which began in early life (see Fig 3 for BMI curves for patients with available growth data); for the remaining 3 patients, the initial symptom was hypernatremia (n = 2; 13%) or polydipsia (n = 1; 6%). The second most common hypothalamic feature was altered water balance for 13 patients (86%), presenting most commonly as hypernatremia (n = 7; 46%) and leading to 6 of those 7 patients being classified as having diabetes insipidus but without a confirmatory water deprivation test. Although the 6 patients who were classified as having diabetes insipidus experienced some improvement with the use of desmopressin, water intake designed to meet minimal...
daily requirements resolved the hypernatremia for all 7 affected patients. Four patients (26%) suffered episodes of hyponatremia (2 classified as transient syndrome of inappropriate antidiuretic hormone secretion, temporarily associated with the use of carbamazepine in 1 patient).

Nine patients (60%) had growth hormone stimulation tests performed; all of them were reported to have a maximal growth hormone response of <10 ng/mL, which is generally considered as evidence for growth hormone deficiency. Although 4 patients displayed deceleration of their growth rate, only 3 of those 4 patients had short stature, defined as height below the 5th percentile. Nevertheless, 6 of the 9 patients were treated with growth hormone, without evidence of improvement in body composition (BMI). With a strict diet and exercise program one of the patients showed significant improvement in BMI, with no evident improvement in clinical features. Other prominent features of HD were hyperprolactinemia (n = 7; 46%), hypothyroidism (n = 5; 33%), adrenal insufficiency (n = 4; 26%), and alterations in pubertal development (n = 4; 26%).

Respiratory Manifestations
Of the 15 patients with available comprehensive medical charts, 9 (60%) experienced cardiorespiratory arrest. For 4 of those 9 patients, there was evidence of abnormal respiratory control before the arrest, manifested as hemoglobin desaturation during sleep for 2, cyanotic episodes during wakefulness for 1, and obstructive sleep apnea for 1. These manifestations were present from 2 months to a few days before the cardiopulmonary arrest. Obstructive sleep apnea was present for 8 of the 15 patients, and it preceded central hypoventilation for 2. All 15 patients demonstrated alveolar hypoventilation; during the same initial evaluation, obstructive sleep apnea was also documented for 5 patients (33%). Seven (47%) of the 15 patients required 24-hour/day artificial ventilation, with the remainder needing support during sleep only. The support was provided with a tracheostomy and mechanical ventilation in 7 cases (47%) and with bilevel positive airway pressure mask ventilation in 8 cases (53%). Patients who required 24-hour/day ventilation had an earlier onset of respiratory manifestations, with a median onset at 3.8 years for the 24-hour/day ventilation group, compared with 7.8 years for the nighttime-only ventilation group (P = .03, Mann-Whitney test). No clinical differences in HD or autonomic dysregulation were found between patients who required continuous versus nighttime-only ventilatory support.

All 9 patients studied in the Respiratory Physiology Laboratory at RUMC were evaluated awake and asleep, in a temperature-controlled room. While awake, all 9 patients demonstrated relative tachypnea during spontaneous breathing, regardless of the measured hemoglo-
<table>
<thead>
<tr>
<th>Phenotypes</th>
<th>Age at Onset, y</th>
<th>Median</th>
<th>25th Percentile</th>
<th>75th Percentile</th>
<th>n (Known Age at Onset)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rapid-onset obesity</td>
<td>3</td>
<td>2.17</td>
<td>4</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Hyperphagia</td>
<td>3.42</td>
<td>2.54</td>
<td>5.17</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Polydipsia</td>
<td>3.58</td>
<td>3.29</td>
<td>5.08</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Polycythemia</td>
<td>3.71</td>
<td>3.54</td>
<td>5.08</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Hyponatremia</td>
<td>5.54</td>
<td>4.65</td>
<td>6.44</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Adrenal insufficiency</td>
<td>6.17</td>
<td>4.79</td>
<td>7.54</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Transient diabetes insipidus</td>
<td>6.17</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Transient SIADH</td>
<td>6.17</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Hypermastria</td>
<td>6.71</td>
<td>2.46</td>
<td>11.79</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Premature puberty</td>
<td>7.21</td>
<td>6.77</td>
<td>7.65</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Diabetes insipid</td>
<td>7.58</td>
<td>2.75</td>
<td>11.83</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Failed growth hormone stimulation test</td>
<td>7.92</td>
<td>5.29</td>
<td>10.54</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Short stature</td>
<td>8.08</td>
<td>6.75</td>
<td>11.29</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Hypothyroidism</td>
<td>8.46</td>
<td>5.31</td>
<td>11.60</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Low IGF-1 and IGFBP-3 levels</td>
<td>8.71</td>
<td>6.31</td>
<td>11.10</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Hyperprolactinemia</td>
<td>8.92</td>
<td>3.92</td>
<td>12.92</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Hypopidipsia</td>
<td>8.96</td>
<td>6.71</td>
<td>11.79</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Hypogonadotropic hypogonadism</td>
<td>11.92</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Amenorrhea</td>
<td>13.17</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Delayed puberty</td>
<td>13.58</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Autonomic dysfunction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Altered sweating</td>
<td>3.58</td>
<td>2.71</td>
<td>6.04</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Gastrointestinal dysmotility</td>
<td>3.83</td>
<td>3.29</td>
<td>5.75</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Tumor of neural crest origin</td>
<td>4.17</td>
<td>2.79</td>
<td>6.83</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Ophthalmic manifestations</td>
<td>4.50</td>
<td>3.21</td>
<td>8.75</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Cold hands and feet</td>
<td>4.54</td>
<td>3.92</td>
<td>7.13</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Thermal dysregulation</td>
<td>7.00</td>
<td>4.46</td>
<td>13.46</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Altered perception of pain</td>
<td>7.00</td>
<td>5.42</td>
<td>8.21</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Bradycardia</td>
<td>7.58</td>
<td>7.13</td>
<td>9.17</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Respiratory manifestations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alveolar hypoventilation</td>
<td>6.17</td>
<td>3.75</td>
<td>8.50</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Cyanotic episodes</td>
<td>6.58</td>
<td>4.75</td>
<td>7.92</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Cardiorespiratory arrest</td>
<td>7.29</td>
<td>3.54</td>
<td>8.25</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Obstructive sleep apnea</td>
<td>7.67</td>
<td>3.42</td>
<td>9.08</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Psychiatric disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>2.17</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Emotional lability</td>
<td>2.17</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Depression</td>
<td>3.33</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Oppositional-defiant disorder</td>
<td>3.50</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Psychosis</td>
<td>4.63</td>
<td>4.02</td>
<td>5.23</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Flat affect</td>
<td>8.50</td>
<td>8.04</td>
<td>8.96</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Developmental disorders</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asperger's syndrome</td>
<td>3.25</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>ADHD</td>
<td>3.25</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Pervasive developmental disorder</td>
<td>3.50</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Developmental regression</td>
<td>3.83</td>
<td>3.46</td>
<td>3.96</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Developmental delay</td>
<td>4.92</td>
<td>3.00</td>
<td>6.83</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Mental retardation</td>
<td>8.50</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Other findings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypersomnolence</td>
<td>2.67</td>
<td>2.08</td>
<td>3.25</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Recurrent pneumonia before ventilatory support</td>
<td>3.29</td>
<td>2.40</td>
<td>4.19</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Seizure</td>
<td>3.92</td>
<td>3.75</td>
<td>3.92</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Hypercholesterolemia</td>
<td>5.08</td>
<td>4.79</td>
<td>10.54</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Deceased</td>
<td>7.58</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Enuresis</td>
<td>7.71</td>
<td>7.44</td>
<td>7.98</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Abnormal brain MRI scans</td>
<td>8.00</td>
<td>7.25</td>
<td>10.42</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Hypotonia</td>
<td>8.17</td>
<td>7.88</td>
<td>8.38</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Scoliosis</td>
<td>8.92</td>
<td>8.83</td>
<td>9.00</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Type 2 diabetes mellitus</td>
<td>11.25</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Impaired glucose tolerance</td>
<td>11.33</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

SIADH indicates syndrome of inappropriate antidiuretic hormone secretion; IGF, insulin-like growth factor; IGFBP, insulin-like growth factor-binding protein; ADHD, attention-deficit/hyperactivity disorder.
bin saturation and end tidal carbon dioxide level (mean ± SD respiratory rate: 34 ± 13 breaths per minute). All 9 patients demonstrated alveolar hypoventilation, with variable severity of resultant hypercarbia (mean ± SD end tidal carbon dioxide level: 56 ± 7 mm Hg) and hypoxemia (mean ± SD hemoglobin saturation: 89 ± 6%) during wakefulness. All 9 children demonstrated alveolar hypoventilation during spontaneous breathing in non-rapid eye movement sleep, but none demonstrated an increase in rate or depth of breathing in response to the resulting endogenously challenges of hypercarbia (mean ± SD end tidal carbon dioxide level: 62 ± 13 mm Hg) and hypoxemia (mean ± SD hemoglobin saturation: 75 ± 17%). In rapid eye movement sleep with artificial ventilatory support, none of the children added “extra breaths” to their mechanical ventilator/bilevel positive airway pressure respiratory rate. One of the 9 children demonstrated obstructive sleep apnea in spontaneous breathing trials during sleep. Two of the 9 patients were reported as being “great swimmers,” with parental reports of cyanosis during swimming.

Among the 9 children studied at RUMC, 5 breathed spontaneously while awake, but 2 of those 5 children required supplemental oxygen because of apparent hypoventilation. Four of the 9 children had a tracheostomy and mechanical ventilation 24 hours/day, typically in the pressure-plateau mode to optimize oxygenation (hemoglobin saturation: ≥95%) and ventilation (end tidal carbon dioxide level: 35–45 mm Hg). Five of the 9 children used bilevel positive airway pressure mask ventilation in the spontaneous timed mode for all sleep time; mask ventilation was unsatisfactory for the 2 children with more-severe awake hypoventilation and oxygen requirements, and a tracheostomy was recommended.

**Autonomic Dysregulation**

Symptoms of autonomic dysregulation were identified for all 15 patients for whom comprehensive medical charts were available. The most common manifestations were ophthalmologic, occurring among a total of 13 patients (86%). Eight patients had pupillary dysfunction (primarily altered responses to light) and 7 strabismus; 4 patients had both. Poor upward gaze and opsonolus were reported for 2 patients each; alacrism, increased blinking, oculomotor apraxia, and ptosis were each reported for 1 patient.

Gastrointestinal dysmotility was reported for 10 patients (66%). Constipation and chronic diarrhea were the most common forms of dysmotility (present for 5 and 4 patients, respectively). Two of the patients with chronic diarrhea developed rectal prolapse; 2 of the patients with constipation were also reported to have gastroesophageal reflux. Thermal dysregulation, manifest as episodes of hyperthermia or hypothermia, were reported for 11 patients (73%). Tumors of neural crest origin were described for 5 (33%) of the 15 patients, that is, ganglioneuroblastoma for 3 and ganglioneuroma for 2. These neoplasms were diagnosed at a median of 2.4 years (range: 0–9 years) after the onset of HD and hypoventilation. Three of those tumors were found in the chest, and 2 were found in the abdomen.

**Developmental Disorders**

Of the 15 patients with comprehensive records, 3 (20%) were reported to have developmental delays, documented before the onset of hypoventilation for 1 patient. One of the patients with developmental delay was later diagnosed as having mild mental retardation. Three other patients (20%) presented developmental regression, 1 before and 2 after the onset of hypoventilation. One of the patients with developmental regression was also diagnosed as having pervasive developmental disorder, attention-deficit/hyperactivity disorder, and Asperger’s syndrome.

**Behavioral Disorders**

Eight (53%) of the 15 patients were reported to have behavioral disorders. Two patients were diagnosed as having depression, 1 of whom was diagnosed as also having Tourette’s syndrome, obsessive-compulsive disorder, and episodes of psychosis.

**Other Findings**

To date, 1 of the 15 patients has died (the patient was found disconnected from ventilatory support, cyanotic with a weak pulse, and could not be resuscitated). Two patients were found to have abnormalities on brain MRI.

![Figure 1](image-url)
scans before suffering cardiorespiratory arrest (a self-resolved Rathke’s cleft cyst and hypointensities in the pons and midbrain). Five patients (33%) were found to have brain MRI abnormalities after experiencing cardiorespiratory arrest (basal ganglia hypointensities for 2 patients and hypointensities in the pons and midbrain, ischemic injury in the frontal, parietal, and occipital lobes, and a partial empty sella for 1 patient each). Five patients (33%) were reported to have generalized tonic-clonic seizures at the time of initial diagnosis or associated with subsequent episodes of hypoxemia. One of the patients has significant seizure activity and no longer communicates verbally.

Other common characteristics are described on Table 1. Abnormalities not included in the aforementioned categories were found in isolated patients and are provided in Tables 1 and 2.

Genetic Testing

**Tested Group**

Genetic testing was performed for 15 of the 23 identified children for the clinical PHOX2B assay and sequencing of the PHOX2B gene was performed on 11 children for the TRKB and BDNF genes. Two samples were studied through high-resolution chromosome analysis, CGH, and subtelomeric FISH.

**PHOX2B**

None of the tested children with the LO-CHS/HD phenotype had a CCHS-related mutation in the PHOX2B gene.

**TRKB and BDNF**

No novel or rare variants were identified in the coding regions of either NTRK2 or BDNF for these patients. Two of the 11 subjects were heterozygous for the previously reported, common BDNF polymorphism Val66Met (rs6265), which reflects the expected frequency in the population.30–33

**High-Resolution Chromosome Analysis, CGH, and Subtelomeric FISH**

High-resolution chromosome analysis, CGH, and subtelomeric FISH results were negative for the 2 patients selected for these analyses.

**Other Testing**

Five (33%) of these patients had previous negative testing for Prader-Willi syndrome (DNA methylation), and 1...
tested negative for the DiGeorge syndrome-associated chromosome 22q11.2 microdeletion. Karyotyping was performed for 4 patients, and results were normal for all of them.

**DISCUSSION**

On the basis of the findings from our systematic analysis of comprehensive medical charts, previously reported cases, and our clear genetic distinction between this syndrome of LO-CHS/HD and CCHS, we propose the term rapid-onset obesity with HD, hypoventilation, and autonomic dysregulation (ROHHAD) for this entity. A remarkable feature of these patients is the apparent normality of their first 2 to 4 years of life, with sudden onset of HD, typically with the onset of rapid weight gain and obesity early in life, followed by autonomic dysregulation and later hypoventilation. There is wide variation in the reported age at onset of autonomic dysfunction, as well as in the interval between the onset of HD and hypoventilation. If it is not identified or is treated inadequately, then the alveolar hypoventilation can be fatal, as evidenced by the high incidence of cardiorespiratory arrest in this group, or induce potential morbidity. The clinical management of these patients requires detailed physiologic assessment, including comprehensive evaluation in the baseline state and with perturbation; evaluation of the hypothalamic-pituitary axis with hormonal replacement when needed; respiratory physiologic assessment during wakefulness and sleep; and MRI or computed tomographic screening of the chest and abdomen for neural crest tumors (ganglioneuromas or ganglioneuroblastomas). Because of the unclear nature of the water balance abnormalities, we recommend performing formal water deprivation tests, with measurements of arginine-vasopressin levels. Brain imaging should be performed to exclude the possibility of hypothalamic-pituitary abnormalities attributable to intracranial lesions. The clinical evaluation should be performed with the understanding that some of the features noted might not be explained by more-common disorders, as illustrated by the fact that many of these patients were classified as having diabetes insipidus, whereas a more appropriate diagnosis might have been hypodipsic hyponatremia and perhaps partial diabetes insipidus. Another example is treatment of these patients with growth hormone on the basis of a failed growth hormone stimulation test. Although abnormalities in the growth hormone axis might exist, these were not clinically evident from the growth patterns; moreover, severe obesity can result in lack of growth hormone responses during stimulation tests.

The available data have allowed us to detail more thoroughly the characteristics of this syndrome, including the earliest presenting symptoms and typical time course, and to document the previously unreported high incidence of cardiorespiratory arrests in this syndrome. Better characterization and the availability of a larger database would be invaluable for advancing knowledge regarding the cause of this syndrome and improving the identification and treatment of these children and might provide key insights into the normal physiologic processes of some of the most basic, vital, neurologic functions. Such characterization could likely guide future candidate gene analysis. Although some features of the CCHS phenotype are seen in patients with ROHHAD, the latter demonstrate an even wider spectrum of involved systems, suggesting a defect in a more-proximal or different genetic pathway involved in autonomic nervous system differentiation or development. The absence of PHOX2B mutations in patients with ROHHAD establishes this syndrome as a separate entity.

On the basis of the phenotype seen in animal studies, BDNF/TRKB signaling seemed to be a reasonable candidate pathway that could contribute to several of the features seen in patients with ROHHAD. Firstly, BDNF-deficient mice have deficits in control of breathing. Indeed, BDNF is known to be important in the develop-

---

**FIGURE 3**

BMI for age in patients with available growth data. A, girls; B, boys. Each line represents one patient. Patients demonstrated rapid-onset obesity in the first 10 years of age. Although rapid weight gain was documented for all patients shown, not all height and weight points were available to allow plotting of some early BMI points.
ment and maintenance of neuronal populations involved in respiratory control. Neurotrophin signaling has also been implicated in the development and function of sensory neurons involved in pain sensation. Homozygous null mutations in Ntrk2, the gene encoding TrkB, are lethal in rodents. However, in the first week of life, surviving Ntrk2−/− mice do not respond to sharp pinpricks in the vibrissae region. Furthermore, TrKB and BDNF have been implicated in the regulation of food intake and body weight. TrkB-hypomorphic mice that expressed 24% of normal levels of TrkB were obese, with increased food intake, as were Bdnf-heterozygous knockout mice and mice with a conditional deletion of Bdnf in the postnatal brain. A loss-of-function mutation in NTRK2 was identified in a patient with severe obesity and hyperphagia who also had impaired nociception. However, mutations in BDNF and NTRK2 were not identified in patients with ROHHAD, which suggests that alternative genes and pathways need to be considered.

High-resolution, conventional, G-banding and subtelomeric FISH are considered standard for the evaluation of individuals with idiopathic mental retardation. Subtelomeric FISH evaluates the gene-rich ends of the chromosomes for rearrangements found in a significant group of patients with mental retardation, with or without additional anomalies. High-resolution metaphase CGH has proved useful for demonstrating interstitial chromosome deletions and duplications as small as 3 megabases. Use of this multimodal approach for 2 patients with a unique disease screens effectively for numerical and cryptic structural chromosome rearrangements within the limits of the technology. Detailed phenotyping and new insights into pathways relevant to the autonomic and endocrine functions of the nervous system should guide us to the genetic basis for this syndrome.

Because of the spectrum of organ systems affected in ROHHAD, the initial medical contact might be a general pediatrician, endocrinologist, pulmonologist, oncologist, or other pediatric subspecialist. If the diagnosis is not considered, then catastrophic consequences may occur, as noted in many of the cases reported. We anticipate that, if a pediatrician or subspecialist notes rapid-onset obesity after 2 years of age and confirms symptoms of autonomic dysregulation, then she or he would refer the patient for comprehensive respiratory physiologic and endocrinologic testing. Similarly, parental observations of cyanosis with prolonged swimming or expertise at breath-holding contests should be heeded and a child with rapid-onset obesity should be evaluated in a pediatric respiratory physiology laboratory as well as by an endocrinologist. Pulmonary physicians treating patients with clinical features consistent with ROHHAD are advised to be aggressive in their respiratory assessment, with comprehensive studies during wakefulness and sleep in a pediatric respiratory physiology laboratory as soon as the diagnosis of ROHHAD is suspected. Once the diagnosis is confirmed, it is essential to monitor the child with serial studies at 3- to 6-month intervals, to ensure optimal oxygenation and ventilation as indicated during wakefulness and sleep, aiming for hemoglobin saturation values of ≥95% and end tidal carbon dioxide values of 35 to 45 mm Hg. With early ventilatory support, highly trained home nursing with continuous pulse oximetry, end tidal carbon dioxide measurements during sleep, and spot checks during wakefulness, and close follow-up monitoring, the ventilatory care for children with ROHHAD can be optimized. Vigilant screening for tumors of neural crest origin should also be a part of ongoing care for children with ROHHAD, with chest and abdominal imaging every 12 to 18 months. If no tumor is identified in 10 years, then it would be reasonable to decrease the frequency of imaging to every 2 years. By providing a detailed clinical description, a name that reflects the main features of this syndrome, and guidelines for management, we aspire to facilitate the earlier recognition, appropriate treatment, and characterization of the molecular origin of this syndrome.

ACKNOWLEDGMENTS
Support was provided by the Scott Robert Conlon Research Fund (Dr Weese-Mayer), the Wellcome Trust (Dr Farooqi), and the Medical Research Council (Ms Gray).

We thank Clifford W. Ragsdale, PhD, and William B. Dobbyns, MD, PhD, from the University of Chicago for their efforts in helping to identify candidate genes. Cynthia Koloboska, RRT, Poutrise Peters, RRT, and Heather Bennett, RRT, for their efforts in compiling the respiratory physiologic data, the physicians who referred these complex cases, and the patients and their parents for providing the medical charts.

REFERENCES
7. Nattie EE, Bartlett D, Rozycki AA. Central alveolar hypoventilation
Screening and Treatment for Lipid Disorders in Children and Adolescents: Systematic Evidence Review for the US Preventive Services Task Force

Elizabeth M. Haney, MDa,b, Laurie Hoyt Huffman, MSa, Christina Bougatsos, BSa, Michele Freeman, MPHa, Robert D. Steiner, MDb, Heidi D. Nelson, MD, MPHa,b,d

aOregon Evidence-based Practice Center, Department of Medical Informatics and Clinical Epidemiology, bDepartment of Medicine, and dDepartments of Pediatrics and Molecular and Medical Genetics, Oregon Health and Science University, Portland, Oregon; dWomen and Children’s Health Research Center, Providence Health System, Portland, Oregon

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

ABSTRACT

OBJECTIVE. This was a systematic evidence review for the US Preventive Services Task Force, intended to synthesize the published evidence regarding the effectiveness of selecting, testing, and managing children and adolescents with dyslipidemia in the course of routine primary care.

METHODS. Literature searches were performed to identify published articles that addressed 10 key questions. The review focused on screening relevant to primary care of children without previously identified dyslipidemias, but included treatment trials of children with dyslipidemia because some drugs have only been tested in that population.

RESULTS. Normal values for lipids for children and adolescents are defined according to population levels (percentiles). Age, gender, and racial differences and temporal trends may alter these statistical cut points. Approximately 40% to 55% of children with elevated total cholesterol and low-density lipoprotein levels will continue to have elevated lipid levels on follow-up. Current screening recommendations based on family history will fail to detect substantial numbers (30%–60%) of children with elevated lipid levels. Drug treatment for dyslipidemia in children has been studied and shown to be effective only for suspected or proven familial monogenic dyslipidemias. Intensive dietary counseling and follow-up can result in improvements in lipid levels, but these results have not been sustained after the cessation of the intervention. The few trials of exercise are of fair-to-poor quality and show little or no improvements in lipid levels for children with monogenic dyslipidemias. Although reported adverse effects were not serious, studies were generally small and not of sufficient duration to determine long-term effects of either short or extended use.

CONCLUSIONS. Several key issues about screening and treatment of dyslipidemia in children and adolescents could not be addressed because of lack of studies, including effectiveness of screening on adult coronary heart disease or lipid outcomes, optimal ages and intervals for screening children, or effects of treatment of childhood lipid levels on adult coronary heart disease outcomes.
Dyslipidemias are disorders of lipoprotein metabolism that result in abnormal excesses of total cholesterol (TC), low-density lipoprotein cholesterol (LDL-C), or triglyceride or deficiency of high-density lipoprotein cholesterol (HDL-C).\textsuperscript{1,2} Dyslipidemia is an established risk factor for coronary heart disease (CHD), which is the leading cause of death for adults in the United States.\textsuperscript{3} Dyslipidemia rarely leads to adverse health outcomes in childhood, but its long-term effects may be considerable. Although no long-term studies of the direct relationship between lipid levels measured in children and CHD later in life have been conducted, this relationship can be inferred. Large epidemiologic studies indicate that children’s lipid levels correlate with those of adult family members.\textsuperscript{4} Children of parents with CHD have a higher prevalence of dyslipidemia in childhood,\textsuperscript{5} and identification of dyslipidemia in children can identify families at increased risk for CHD.\textsuperscript{6} Studies of children and young adults who died accidentally have reported correlations between lipid levels and arterial fat deposition\textsuperscript{6,7} and noted early lesions of atherosclerosis (fatty streaks) in the abdominal aorta at 3 years of age, coronary arteries at 10 years of age, and further progression with age.\textsuperscript{8–12} Increasing prevalence of risk factors for CHD among children, including metabolic syndrome and obesity, as well as continued emphasis on primary prevention of CHD has raised interest in screening children for dyslipidemia.\textsuperscript{13–15}

Dyslipidemia is defined by laboratory testing and statistically determined criteria. An elevated LDL-C level is the most common clinically significant marker of dyslipidemia in children. The majority of children with dyslipidemia will have idiopathic dyslipidemias (polygenic, risk factor–associated, or multifactorial), whereas a minority will have monogenic or secondary dyslipidemias. The more common genetic dyslipidemias include familial hypercholesterolemia (FH), familial combined hyperlipidemia (FCH), familial defective apoprotein-B, and familial hypertriglyceridemia. Most treatment recommendations advise a low-fat, low-cholesterol diet, such as the American Heart Association (AHA) Step I diet, for children with dyslipidemia beginning at the age of 2 years or older.\textsuperscript{14} Children younger than 2 years should not be prescribed a low-fat, low-cholesterol diet, because their rapid growth and development require adequate fat and cholesterol intake.\textsuperscript{16,17} Children and adolescents with FH or FCH are the only nonadults for whom trials of drug therapy are available and drugs are approved by the US Food and Drug Administration. Bile-acid–binding resins are the only medications approved for treatment of dyslipidemia for children younger than 8 years of age. 3-Hydroxy-3-methylglutaryl coenzyme A (HMG Co-A) reductase inhibitors (statins) are approved for use in older children with heterozygous FH.\textsuperscript{18,19} Other medications used in adults for treatment of hyperlipidemia, such as niacin, are either not recommended for children or have not been adequately evaluated for safety and efficacy in children. Additional interventions for children include dietary supplements (fiber, sterol or stanol margarines, and omega-3 fatty acids), exercise, weight loss for overweight children, and identification and treatment of diabetes mellitus or other causes of secondary dyslipidemia.

The relationship between childhood and adult dyslipidemia, increasing prevalence of related CHD risk factors in children (eg, obesity and diabetes),\textsuperscript{13–15} and continued emphasis on a primary prevention approach for CHD has raised interest in screening children for dyslipidemia. Identifying children with dyslipidemia could lead to interventions or treatments that could prevent or delay adult dyslipidemia and CHD. This rationale lends support to consideration of screening for dyslipidemia as part of well-child care and at other opportunities. Clinic-based screening, neonatal screening, community-based screening, and other prevention strategies have been proposed, but most recommendations support selective strategies to test children who have family members with dyslipidemia or premature CHD and those with unknown family histories.\textsuperscript{16–20}

This evidence review focuses on the strengths and limitations of evidence for identifying and managing children and adolescents with dyslipidemia determined by screening in the course of routine primary care. Our objective was to determine the balance of potential benefits and adverse effects of screening for development of guidelines by the US Preventive Services Task Force (USPSTF). The target population includes children and adolescents 0 to 21 years old without previously known conditions associated with dyslipidemia. There is potential to identify children and adolescents with dyslipidemia in this population from among 3 groups: those with undiagnosed monogenic dyslipidemias such as FH; those with undiagnosed secondary causes of dyslipidemia (diabetes, nephrotic syndrome, hypothyroidism, others); and those with idiopathic dyslipidemia (polygenic, risk factor–associated, or multifactorial) (Fig 1). Although children and adolescents with idiopathic dyslipidemia generally have less severe lipid-level abnormalities than children and adolescents with monogenic disorders, such abnormal levels could still potentially improve with intervention.

METHODS

Evidence reviews for the USPSTF follow a specific methodology\textsuperscript{21} (Fig 2). Key questions examine a chain of evidence about the accuracy and feasibility of screening children and adolescents for dyslipidemia in primary care or community settings (key question 1), abnormal lipid values (key question 2a), appropriate tests (key question 2b), tracking of lipid levels through childhood to adulthood (key question 2c), accuracy of family his-
tory (key question 2d), role of risk factors in selecting children and adolescents for screening (key question 2e), effectiveness of interventions for children and adolescents identified with dyslipidemia (key questions 4–8 and 10), and adverse effects of screening and interventions (key questions 3 and 9).

Studies that addressed key question 1 (Fig 2) include all components in the continuum of the screening process: the screening evaluation, diagnostic evaluation for those identified by the screening results, interventions for those diagnosed with dyslipidemia, and outcome measures that allow determination of the effectiveness of the overall screening process.

Studies of children with previously diagnosed conditions that are known to cause dyslipidemia were not included, because the scope of this review is screening children without known diagnoses. Specifically, studies of children with diabetes were not included, because these children would already be under surveillance for dyslipidemia as a result of their primary disease. This review includes treatment trials of children and adolescents who used dietary, exercise, and drug interventions. Trials of drug therapy in children with heterozygous FH or FCH are included, because drug-treatment trials have been conducted exclusively in this population.

Relevant studies were identified from multiple searches of Medline (1966 through September 2005).22 We obtained additional articles from recent systematic reviews, reference lists of related studies, reviews, editorials, and Web sites and from consulting experts. Retrieved abstracts were entered into an electronic database (EndNote; Thomson ResearchSoft, Carlsbad, CA).

Investigators reviewed all identified abstracts and determined eligibility by applying inclusion and exclusion criteria specific to each key question. Full-text articles of included abstracts were reviewed for relevance. Eligible studies were English language and applicable to US clinical practice and provided primary data relevant to the key questions. Studies of risk factors were included only if they provided multivariate adjusted analyses.

For treatment studies, full-text randomized, controlled trials (RCTs), noncontrolled clinical trials, and noncontrolled prospective studies that provided data on the treatment of children and adolescents with drug therapy, diet, exercise, or combinations of these interventions were reviewed initially. Subsequently, only RCTs and meta-analyses of RCTs that reported serum
lipid outcomes were included. Crossover trials were included if they reported data before crossover. For key question 10, outcomes included either adult lipid levels or adult CHD. Information about adverse effects of treatment was obtained from RCTs and additional sources such as nonrandomized, controlled treatment trials and noncomparative studies of treatment.

Data were extracted from each study, entered directly into evidence tables, and summarized. Benefits and adverse effects of therapies were considered equally important, and both types of outcomes were abstracted. Trials of therapy for children and adolescents with dyslipidemia were categorized by population and intervention. Two reviewers independently rated the RCTs’ quality by using USPSTF criteria21 (Appendix).

**RESULTS**

Our literature search identified 2507 unique citations including 144 articles about screening and testing for dyslipidemia (key question 2); 43 about interventions and tracking of lipid values over time (key questions 4–8 and 10); 6 about the adverse effects of screening (key question 3); and 84 about adverse effects of treatment (key question 9).

**Key Question 1:** Is Screening for Dyslipidemia in Children/Adolescents Effective in Delaying the Onset and Reducing the Incidence of CHD-Related Events?

No studies evaluated the effect of screening children and adolescents on adult lipid-level or disease outcomes.

**Key Question 2:** What Is the Accuracy of Screening for Dyslipidemia in Identifying Children/Adolescents at Increased Risk of CHD-Related Events and Other Outcomes?

**Key Question 2a:** What Are Abnormal Lipid Values in Children/Adolescents?

Although several studies conducted in the United States during the 1970s obtained lipid levels from large samples of normal healthy children,23–25 current recommendations14,16,20,26 are based on distributions of lipid and lipoprotein levels obtained from the Lipid Research Clin-
ics (LRC) Prevalence Study. This study included 1 Canadian and 9 US sites and enrolled subjects primarily on the basis of residency within census tracts, school enrollment, and employment in occupational and industrial groups. Fasting (≥12 hours) lipoprotein levels were obtained from 15,626 children 0 to 19 years old between 1972 and 1976. The selected populations included a broad range of geographic, socioeconomic, occupational, gender, and ethnic groups but were not selected to be a representative sample of the North American population.

In the LRC sample, TC levels increased from birth and stabilized at approximately 2 years of age. At puberty, TC levels declined slightly for both boys and girls, and HDL-C levels declined for boys. For all children, the mean serum level for TC was ~160 mg/dL and for LDL-C was 100 mg/dL. The 95th percentile level was 200 mg/dL for TC and 130 mg/dL for LDL-C. Although the results for black children were similar, they were based on smaller numbers and provided only TC and triglyceride data.

More recent data from the National Health and Nutrition Examination Survey III (1988–1994) were derived from 7499 children and adolescents aged 4 to 19 years. These data provided 95th-percentile levels of 216 mg/dL for serum TC and 152 mg/dL for LDL-C.28 Mean age-specific TC levels peaked at 171 mg/dL at 9 to 11 years and declined at older ages. Girls had significantly higher mean TC and LDL-C levels than boys (P<.005). Non-Hispanic black children and adolescents had significantly higher mean TC, LDL-C, and HDL-C levels compared with non-Hispanic white and Mexican-American children and adolescents. In linear regression models of these data, age, gender, and race have significant effects on lipid levels, which raises questions about the utility of fixed screening cut points.

**Key Question 2b: What Are the Appropriate Tests? How Well Do Screening Tests (Nonfasting TC, Fasting TC, Fasting Lipoprotein Analysis) Identify Children and Adolescents With Dyslipidemia?**

In the American Academy of Pediatrics (AAP) and National Cholesterol Education Program (NCEP) guidelines, TC is used as an initial laboratory measurement for children tested because of a family history of high cholesterol or vascular disease, and a lipoprotein profile is obtained if the patient has a TC over a certain defined target.16,20 In children, the LDL-C level is the basis for initiating treatment and determining goals of therapy.

How well TC levels detect elevated LDL-C levels has been examined with LRC data (ages 6–19, n = 1325)10 and data from the biracial Bogalusa cohort (ages 5–17, n = 2857).31 Elevated levels were defined as >95th percentile. With LRC data, an elevated fasting TC level identified children with elevated LDL-C and triglyceride levels with 69% sensitivity and 98% specificity.30 In the Bogalusa cohort, elevated TC levels detected elevated LDL-C levels with 44% (white females) to 50% (white males, black males and females) sensitivity and 90% specificity (black and white males and females).

In adults, both TC and HDL-C levels are recommended for screening. Although this has not been recommended in guidelines for children and adolescents, it is common in practice (E. Neufeld, MD, PhD [Boston, MA], personal communication regarding screening tests for children, 2005). HDL-C may help distinguish false-negative from true-negative results when used with TC.30 In 260 black adolescents aged 12 to 20 years, fasting TC minus HDL-C above the 95th percentile was 88% to 96% sensitive and 98% specific for predicting an LDL-C level of ≥130 mg/dL.32 Using a lower threshold of fasting TC (≥75th percentile) to detect LDL-C levels ≥95th percentile in a sample of Hispanic children aged 4 to 5, sensitivities were 86% (using an LRC-defined 75th percentile) and 96% (using the sample-defined 75th percentile), and specificities were 93% (LRC defined) and 87% (sample defined).33 A TC level of >215 mg/dL is required, however, to accurately identify a child with elevated LDL-C levels with 95% confidence. No single TC value places a child in the borderline category (170–200 mg/dL) with 95% confidence.34 Direct measurement of LDL-C levels can be made by using nonfasting serum samples and may be as precise as calculated LDL-C levels, but this remains controversial.35,36

**Key Question 2c: How Well Do Lipid Levels Track From Childhood to Adulthood?**

Twenty-three prospective cohort studies contributed information on tracking lipid levels during childhood.37–59 These studies drew from 7 US cohorts and 8 non-US cohorts. Approximately 40% to 55% of children with elevated lipid levels, defined by percentile within a population distribution, will continue to have elevated lipid levels on follow-up (4–15 years later).22 None of these studies, however, evaluated the proportion of children and adolescents with lipid levels >95th percentile who remained in the top 5% at follow-up.

**Key Question 2d: What Is the Accuracy of Family History in Determining Risk?**

Several good-quality studies of diagnostic accuracy evaluated the sensitivity and specificity of family-history information in determining risk for dyslipidemia in children and adolescents (Table 1).32,33,60–73 Studies used different definitions of family history, such as any parental history of heart attack, other parental risk factors, and varying age definitions of early CHD, and selected different levels of LDL-C or TC as the lipid-detection threshold. For example, parental history of early CHD alone was 5% to 17% sensitive for TC >170 mg/dL or LDL-C >130 mg/dL,33,62 whereas parental or grandpar-
<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>Population</th>
<th>Method</th>
<th>Thresholda</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>No. Eligible for Screening (on the Basis of Population of 1000)b</th>
<th>No. Missed (on the Basis of Population of 1000)b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bell and Joseph (1990)</td>
<td>1140 5th-graders</td>
<td>Family history of high cholesterol or MI at &lt;50 y of age in parent or grandparent</td>
<td>Nonfasting TC &gt; 200 mg/dL</td>
<td>64</td>
<td>47</td>
<td>540</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>1140 5th-graders</td>
<td>As stated above, plus family history of stroke, angina, or hypertension</td>
<td>Nonfasting TC &gt; 200 mg/dL</td>
<td>77</td>
<td>24</td>
<td>760</td>
<td>31</td>
</tr>
<tr>
<td>Davidson et al (1991)</td>
<td>1118 4th-graders</td>
<td>Family history from parents (regarding parents, siblings, grandparents, aunts, uncles); early MI defined as that at &lt;56 y of age for men and women</td>
<td>TC &gt; 200 mg/dL</td>
<td>41</td>
<td>68</td>
<td>330</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>1118 4th-graders</td>
<td>Parental questionnaire, definition using AAP criteria for early CHD (&lt;50 y for men, &lt;60 y for women)</td>
<td>TC &gt; 200 mg/dL</td>
<td>31</td>
<td>66</td>
<td>330</td>
<td>96</td>
</tr>
<tr>
<td>Dennison et al (1989), Bogalusa Heart Study</td>
<td>1214 4–10 y</td>
<td>Parental questionnaire asking parental history of any vascular disease (CHD, HTN, diabetes, stroke)</td>
<td>fasting TC ≥ 95th percentile</td>
<td>38 (W), 27 (B)</td>
<td>73 (W), 65 (B)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>2099 11–17 y</td>
<td>Parental questionnaire asking parental history of any vascular disease (CHD, HTN, diabetes, stroke)</td>
<td>fasting TC ≥ 95th percentile</td>
<td>59 (W), 25 (B)</td>
<td>67 (W), 56 (B)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>1214 4–10 y</td>
<td>Parental questionnaire asking parental history of any vascular disease (CHD, HTN, diabetes, stroke)</td>
<td>fasting LDL ≥ 95th percentile</td>
<td>41 (W), 20 (B)</td>
<td>73 (W), 63 (B)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>2099 11–17 y</td>
<td>Parental questionnaire asking parental history of any vascular disease (CHD, HTN, diabetes, stroke)</td>
<td>fasting LDL ≥ 95th percentile</td>
<td>37 (W), 22 (B)</td>
<td>67 (W), 56 (B)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Diller et al (1995), Cincinnati MI Hormone Study</td>
<td>232 2–19 y</td>
<td>Parental questionnaire using NCEP definition of family history of premature CVD</td>
<td>LDL ≥ 130 mg/dL</td>
<td>17</td>
<td>75</td>
<td>246</td>
<td>207</td>
</tr>
<tr>
<td></td>
<td>232 2–19 y</td>
<td>Parental questionnaire asking family history of cholesterol level ≥ 240 mg/dL</td>
<td>LDL ≥ 130 mg/dL</td>
<td>61</td>
<td>74</td>
<td>293</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>232 2–19 y</td>
<td>Both family history of elevated cholesterol level and premature CVD</td>
<td>LDL ≥ 130 mg/dL</td>
<td>74</td>
<td>55</td>
<td>478</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>232 2–19 y</td>
<td>Other indicators: obesity, smoking, use of lipid-raising medications, high-fat diet, HTN</td>
<td>LDL ≥ 130 mg/dL</td>
<td>17.4 (obesity), 9–48 (others)</td>
<td>86 (obesity); 69–95 (others)</td>
<td>547</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>232 2–19 y</td>
<td>Family history of premature CHD (NCEP definition), TC ≥ 240 mg/dL, or any other risk factor (obesity, smoking, lipid-raising medication, high-fat diet, or HTN)</td>
<td>LDL ≥ 130 mg/dL</td>
<td>96</td>
<td>28</td>
<td>746</td>
<td>13</td>
</tr>
<tr>
<td>Gagliano et al (1993)</td>
<td>224 11–20 y</td>
<td>Family history of early MI (&lt;50 y for men, &lt;60 y for women) or elevated lipid levels (TC &gt; 200 mg/dL), history obtained from adolescent</td>
<td>TC &gt; 85th percentile for gender</td>
<td>36</td>
<td>69</td>
<td>320</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>224 11–20 y</td>
<td>Family history as stated above, history obtained from parent</td>
<td>TC &gt; 85th percentile for gender</td>
<td>65</td>
<td>46</td>
<td>589</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>224 11–20 y</td>
<td>Use of combined family history from adolescent and parent</td>
<td>TC &gt; 85th percentile for gender</td>
<td>45</td>
<td>69</td>
<td>361</td>
<td>80</td>
</tr>
<tr>
<td>Author (Year)</td>
<td>Population</td>
<td>Method</td>
<td>Threshold*</td>
<td>Sensitivity, %</td>
<td>Specificity, %</td>
<td>No. Eligible for Screening (on the Basis of Population of 1000)*</td>
<td>No. Missed (on the Basis of Population of 1000)*</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
<td>--------</td>
<td>------------</td>
<td>----------------</td>
<td>----------------</td>
<td>---------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Griffin et al (1989), 8 office practices</td>
<td>1005 2–13 y</td>
<td>Parental and grandparental history of hypercholesterolemia or CHD at &lt;55 y</td>
<td>Fasting LDL &gt;95th percentile</td>
<td>46</td>
<td>NR</td>
<td>NA</td>
<td>147</td>
</tr>
<tr>
<td></td>
<td>1005 2–13 y</td>
<td>Parental and grandparental history of any risk factor or complication (hypercholesterolemia, diabetes, HTN, gout, obesity, or atherosclerosis before age 55 y)</td>
<td>Fasting LDL &gt;95th percentile</td>
<td>78</td>
<td>NR</td>
<td>NA</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>1005 2–13 y</td>
<td>Parental and grandparental history of hypercholesterolemia or CHD at &lt;55 y</td>
<td>Fasting LDL &gt;90th percentile</td>
<td>51</td>
<td>63</td>
<td>385</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>1005 2–13 y</td>
<td>Any history of parent or grandparent with a risk factor or complication (hypercholesterolemia, diabetes, HTN, gout, obesity, or atherosclerosis before age 55 y)</td>
<td>Fasting LDL &gt;90th percentile</td>
<td>51</td>
<td>63</td>
<td>385</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>1005 2–13 y</td>
<td>Overweight (weight for height &gt;95th percentile) plus family history of early CHD or hypercholesterolemia</td>
<td>Fasting LDL &gt;95th percentile</td>
<td>46</td>
<td>NR</td>
<td>NA</td>
<td>147</td>
</tr>
<tr>
<td></td>
<td>1005 2–13 y</td>
<td>Overweight (weight for height &gt;95th percentile) plus family history of any risk factor or complication</td>
<td>Fasting LDL &gt;95th percentile</td>
<td>78</td>
<td>NR</td>
<td>NA</td>
<td>59</td>
</tr>
<tr>
<td>Muhonen et al (1994), Muscatine, IA</td>
<td>599 14–20 y</td>
<td>Parental history of high cholesterol</td>
<td>Highest decile of fasting TC</td>
<td>34</td>
<td>76</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>599 14–20 y</td>
<td>Parental history of high cholesterol</td>
<td>Highest decile of fasting LDL</td>
<td>34</td>
<td>76</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>599 14–20 y</td>
<td>Parental history of high cholesterol</td>
<td>Lowest decile of fasting HDL</td>
<td>26</td>
<td>75</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>O’Loughlin et al (2004), Quebec</td>
<td>2217 9, 13, and 16 y</td>
<td>Parental questionnaire asking personal history of (1) high cholesterol, (2) medications for cholesterol, (3) MI or angina, (4) stroke, CVD, or PVD, or (5) medications for the heart; unknown family history coded as negative</td>
<td>Fasting LDL ≥ 109 mg/dL (&quot;borderline&quot;)</td>
<td>33</td>
<td>76</td>
<td>256</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>2217 9, 13, and 16 y</td>
<td>Parental questionnaire asking personal history of (1) high cholesterol, (2) medications for cholesterol, (3) MI or angina, (4) stroke, CVD, or PVD, or (5) medications for the heart; unknown family history coded as negative</td>
<td>Fasting LDL ≥ 131.5 mg/dL (&quot;high&quot;)</td>
<td>41</td>
<td>75</td>
<td>256</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>2217 9, 13, and 16 y</td>
<td>Parental questionnaire asking personal history of (1) high cholesterol, (2) medications for cholesterol, (3) MI or angina, (4) stroke, CVD, or PVD, or (5) medications for the heart; unknown family history excluded</td>
<td>Fasting LDL ≥ 109 mg/dL (&quot;borderline&quot;)</td>
<td>42</td>
<td>70</td>
<td>NA</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>2217 9, 13, and 16 y</td>
<td>Parental questionnaire asking personal history of (1) high cholesterol, (2) medications for cholesterol, (3) MI or angina, (4) stroke, CVD, or PVD, or (5) medications for the heart; unknown family history excluded</td>
<td>Fasting LDL ≥ 131.5 mg/dL (&quot;high&quot;)</td>
<td>51</td>
<td>69</td>
<td>NA</td>
<td>19</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
<td>--------</td>
<td>-----------</td>
<td>---------------</td>
<td>---------------</td>
<td>---------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Primrose et al (1994), Ireland</td>
<td>1012</td>
<td>12–15 y</td>
<td>History of stroke, angina, or MI in either parent at any age or in first-degree grandparents, uncles, or aunts at &lt;55 y; questionnaires completed by parents</td>
<td>Nonfasting TC &gt;95th percentile according to LRC</td>
<td>33</td>
<td>72</td>
<td>293</td>
</tr>
<tr>
<td>Resnicow et al (1993)</td>
<td>574</td>
<td>Elementary school age</td>
<td>Parental cholesterol ≥240 mg/dL in 1 parent only with known and reported value by that parent</td>
<td>Nonfasting TC &gt;200 mg/dL</td>
<td>10</td>
<td>91</td>
<td>90</td>
</tr>
<tr>
<td>Rifai et al (1996)</td>
<td>260</td>
<td>12–20 y</td>
<td>Family history of early CHD or hyperlipidemia</td>
<td>Fasting LDL &gt;110 mg/dL</td>
<td>10</td>
<td>NR</td>
<td>365</td>
</tr>
<tr>
<td>Sanchez Bayle et al (1992), Spain</td>
<td>2224</td>
<td>2–18 y</td>
<td>Parental history of MI</td>
<td>Fasting TC &gt;200 mg/dL</td>
<td>7</td>
<td>96</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>2224</td>
<td>2–18 y</td>
<td>Parental history of MI</td>
<td>Fasting LDL &gt;135 mg/dL</td>
<td>9</td>
<td>96</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>2224</td>
<td>2–18 y</td>
<td>Parental history of stroke, HTN, diabetes, or hypercholesterolemia (but not MI)</td>
<td>Fasting TC &gt;200 mg/dL</td>
<td>14</td>
<td>90</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>2224</td>
<td>2–18 y</td>
<td>Parental history of stroke, HTN, diabetes, or hypercholesterolemia (but not MI)</td>
<td>Fasting LDL &gt;135 mg/dL</td>
<td>14</td>
<td>91</td>
<td>98</td>
</tr>
<tr>
<td>Shea et al (1990), Study of Childhood Activity and Nutrition</td>
<td>108 Hispanic</td>
<td>4–5 y</td>
<td>AAP definition (maternal hypertension, diabetes, obesity, hyperlipidemia, or family history of premature CHD or hyperlipidemia)</td>
<td>Fasting TC &gt;170 mg/dL</td>
<td>57</td>
<td>59</td>
<td>493</td>
</tr>
<tr>
<td></td>
<td>108 Hispanic</td>
<td>4–5 y</td>
<td>AHA and NIH Consensus Conference definition (history of hyperlipidemia or premature CHD in the child's parent, aunt, uncle, or grandparent)</td>
<td>Fasting TC &gt;170 mg/dL</td>
<td>46</td>
<td>70</td>
<td>352</td>
</tr>
<tr>
<td></td>
<td>108 Hispanic</td>
<td>4–5 y</td>
<td>NCEP guidelines (history of MI or sudden death in the child's parent, aunt, uncle, or grandparent; CHD before age 55 y)</td>
<td>Fasting TC &gt;170 mg/dL</td>
<td>5</td>
<td>92</td>
<td>74</td>
</tr>
<tr>
<td>Steiner et al (1991), Kaiser population</td>
<td>1001 (88% Hispanic, 33.5% W, 15% B, 11% Asian)</td>
<td>12–21 y</td>
<td>AAP 1998 criteria (known hyperlipidemia in parent or sibling, known MI/angina, current corticosteroid use, juvenile diabetes, hypothyroidism, renal/endoctrine/hepatic disease in teenager)</td>
<td>Nonfasting TC ≥200 mg/dL, repeated fasting TC if initial result was ≥200 mg/dL, repeated a third time if &gt;30 mg/dL, variability between the first 2 measurements</td>
<td>63</td>
<td>60</td>
<td>400</td>
</tr>
<tr>
<td>Troxler et al (1991)</td>
<td>110, mostly Hispanic</td>
<td>Senior high school student</td>
<td>Questionnaires completed with parental assistance; family history in parents or grandparents of high cholesterol or CHD at &lt;55 y (AAP)</td>
<td>Fasting TC &gt;75th percentile (175 mg/dL)</td>
<td>38</td>
<td>79</td>
<td>218</td>
</tr>
<tr>
<td>Wadowski et al (1994)</td>
<td>300</td>
<td>2–14 y</td>
<td>Family history of CHD in parent or grandparent at &lt;55 y (AAP)</td>
<td>Fasting TC &gt;215 mg/dL</td>
<td>59</td>
<td>72</td>
<td>327</td>
</tr>
</tbody>
</table>

MI indicates myocardial infarction; HTN, hypertension; CVD, cardiovascular disease; PVD, peripheral vascular disease; NA, not applicable; NIH, National Institutes of Health; W, white; B, black.

a If not explicitly stated, values are mixed nonfasting/fasting or not reported.
b Number eligible for screening and number missed were calculated from available data. In some cases, reported data did not allow for these calculations (indicated with NA).
ental history of early CHD was 46% sensitive for LDL > 95th percentile.\(^6\)

Regardless of the precise definition, using positive family-history information to trigger lipid testing misses substantial numbers of children with elevated lipid levels, ranging from 17% to 90% overall and 30% to 60% in most studies.\(^12,63,64,67,69,71,74-76\) The proportion of children and adolescents who qualify for screening on the basis of family history is generally between 25% and 55% depending on the sensitivity of the specific family-history question.\(^8\)

**Key Question 2c: What Are Other Important Risk Factors?**
Forty-three cohort and cross-sectional studies of mixed quality with adjusted statistical analyses contributed information on additional risk factors for identifying children at increased risk for elevated lipid levels and/or CHD-related events.\(^5,5,78-119\) Thirty studies examined overweight or body fat composition measures as a risk factor for dyslipidemia.\(^9\) These measures were the most consistently effective in predicting risk of dyslipidemia compared with other factors assessed.\(^22\) Childhood overweight, as measured by BMI, was the best independent predictor of adult dyslipidemia after LDL-C level, specifically when considering BMI increases from childhood to adulthood.\(^120\) Of 6 studies that evaluated overweight as a risk, 5 found that overweight was associated with abnormal lipid levels.\(^84,85,93,109,114,116\)

**Key Question 2f: What Are Effective Screening Strategies for Children/Adolescents (Including Frequency of Testing, Optimal Age for Testing)?**
Thirty-two studies evaluated screening strategies among children in various settings.\(^\dagger\) The only RCT compared 2 regimens for screening college students.\(^130\) All others were noncomparative prospective studies that described screening interventions and differed considerably in venue (school, pediatric clinic, hospital, or population-based cohort), methods (fasting or nonfasting samples, method for detecting positive family history), and outcomes. Most of them reported low parental compliance with follow-up testing\(^75,135-138\) even when follow-up was provided free of charge, as in prepaid health plans.

Studies demonstrated low compliance among primary care physicians in following current guidelines for screening.\(^139\) In an ancillary study of the Child Adolescent Trial for Cardiovascular Health (CATCH), parents were given recommendations to consult their child’s physician if his or her TC level exceeded 200 mg/dL on ≥1 occasion.\(^140\) After physicians examined the children, only 59% were evaluated further for possible elevated cholesterol levels. Of these, half of the physicians repeated cholesterol tests, 42% asked about family history, 38% made recommendations for dietary management, and only 12% referred children to dietitians.\(^140\)

Neonatal screening for dyslipidemia has been examined in multiple studies of cord blood testing.\(^13,141-159\) dried filter paper blood spots from cord blood,\(^155\) or heel sticks of 3- to 7-day-old infants.\(^156,161\) No studies screened a general population of infants and followed abnormal results with mutation analysis or LDL-C receptor activity assays, which makes it difficult to determine the value of such screening.

**Key Question 3: What Are the Adverse Effects of Screening (Including False-Positive and False-Negative Results, Labeling, etc)?**
Potential adverse effects of screening for dyslipidemia among children were examined in 1 RCT\(^162\) and 5 noncomparative studies.\(^79,135-138\) Although 1 small study showed increased parental reporting of behavior difficulties among children with dyslipidemia, these reports were not confirmed objectively.\(^138\) No studies reported increased anxiety or depression among screened children or their parents.\(^136-138\)

**Key Question 4: In Children/Adolescents, What Is the Effectiveness of Drug, Diet, Exercise, and Combination Therapy in Reducing the Incidence of Adult Dyslipidemia and Delaying the Onset and Reducing the Incidence of CHD-Related Events (Including Optimal Age for Initiation of Treatment)?**
No studies evaluated the effect of a childhood intervention on the incidence of adult dyslipidemia or CHD-related events and outcomes.

**Key Questions 5–8: What Is the Effectiveness of Drug, Diet, Exercise, and Combination Therapy for Treating Dyslipidemia in Children/Adolescents?**
Forty RCTs that met the inclusion criteria addressed the effectiveness of interventions for treatment of dyslipidemia in children and adolescents.\(^18,19,163-200\) Statins, bile-acid–binding resins, and fibrates have been tested and reported only in children with FH and FCH. Applicability of results from these trials to children without these conditions may be limited. In addition, 18 studies used populations recruited from single lipid clinics.\(\S\) Major limitations of trials include < 20 subjects in each study arm,\(\|\) high loss to follow-up,\(\|\) failure of blinding,\(\|\) lack of results presented for the period before crossover,\(\#\) lack of intention-to-treat analyses,\(\#\) and lack of data reported for the placebo group.\(\#\)

**Studies in Children With Probable or Definite FH**

**Drug Treatment**
Eleven trials evaluated drug therapies for treatment of children with probable or definite heterozygous FH (Ta-
Most of these studies included children who were already compliant with a recommended low-saturated-fat, low-cholesterol diet, and both treatment and control groups were maintained on the diet during the trials. All the trials of statin drugs†† demonstrated improvement in TC and LDL-C levels among children and adolescents with FH. The decrease in TC compared with baseline ranged from 17% to 32% for subjects in the treatment groups versus changes of /H110013.6% to /H110022.3% for those in the placebo groups. The decreases in LDL-C level ranged from 19% to 41% for subjects in the treatment groups versus changes of /H110010.67% to /H110023% for those in the placebo groups. Changes in HDL-C and triglyceride levels were mixed.‡‡

Trials of cholestyramine186 and colestipol185 demonstrated decreased TC and LDL-C levels but no change in HDL-C or triglyceride levels. Trials that evaluated bezafibrate,192 vitamins C and E,181 docosahexaenoic acid,198,200 p-aminosalicylic acid,184 combined colestipol and pravastatin versus colestipol alone,165 and powder versus pill form of cholestyramine173 failed to report precross-over data.

Diet Treatment

Five trials that evaluated diet treatments in children with FH or FCH met inclusion criteria.166,167,177,179,199 Although trials of sterol margarines and psyllium were crossover trials without precrossover results presented, the wash-out periods between treatment phases were 4 to 6 weeks, suggesting that results may be valid.166,177,179 Reductions in TC and LDL-C levels were significant in these trials (reduction of 7.4%–11% and 10%–14%, respectively). There was no significant improvement in lipid levels with 8 weeks of treatment with garlic extract.199

Exercise Treatment

No studies evaluated exercise treatment for lowering lipid levels in children with FH.

Studies in Children With Elevated Lipid Levels but Not Meeting Criteria for FH

Drug Treatment

No studies evaluated drug interventions in children without monogenic dyslipidemia.

Diet Treatment

Dietary interventions in general populations of children and adolescents were addressed in 7 studies (Table 3).§§ A trial conducted by the Dietary Intervention Study in Children (DISC) Collaborative Research Group showed that intensive dietary counseling over 3 years was effective (8% improvement in LDL-C level compared with control)170 but not sustained at 5- and 7-year follow-ups once the intervention ceased.169 A study of the Parent-Child AutoTutorial (PCAT) program173 reported 8% improvement in LDL-C level compared with the at-risk control group (P < .05). One trial of psyllium did not present precrossover data.80

### TABLE 2  RCTs of Drug Treatment for Children With Monogenic Dyslipidemia

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>Drug</th>
<th>Population</th>
<th>Duration of Trial</th>
<th>Significant Changes vs Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>Age, y</td>
<td>TC</td>
</tr>
<tr>
<td>Statins</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Claussen et al19 (2005)</td>
<td>Lovastatin 20 vs 40 mg/d vs placebo</td>
<td>54 girls</td>
<td>11–18</td>
<td>24 wk</td>
</tr>
<tr>
<td>Couture et al18 (1998)</td>
<td>Simvastatin 20 mg/d vs placebo</td>
<td>63</td>
<td>8–17</td>
<td>6 wk</td>
</tr>
<tr>
<td>de Jongh et al104 (2002)</td>
<td>Simvastatin 10 mg/d, doubled every 8 wk up to 40 mg/d vs placebo</td>
<td>50</td>
<td>9–18</td>
<td>28 wk</td>
</tr>
<tr>
<td>de Jongh et al102 (2002)</td>
<td>Simvastatin 10 mg/d titrating up to 40 mg/d vs placebo</td>
<td>173</td>
<td>10–17</td>
<td>48 wk</td>
</tr>
<tr>
<td>Knipscheer et al112,113 (1996)</td>
<td>Pravastatin in 3 active drug groups, 5, 10, or 20 mg/d, vs placebo</td>
<td>72</td>
<td>11–17</td>
<td>12 wk</td>
</tr>
<tr>
<td>Lambert et al183 (1996)</td>
<td>Lovastatin at 10, 20, 30, or 40 mg/d (4 active drug groups, no placebo)</td>
<td>69 boys</td>
<td>≤ 17</td>
<td>8 wk</td>
</tr>
<tr>
<td>McCrindle et al184 (2003)</td>
<td>Atorvastatin 10 mg/d vs placebo</td>
<td>187</td>
<td>10–17</td>
<td>26 wk</td>
</tr>
<tr>
<td>Stein et al185 (1999)</td>
<td>Lovastatin starting at 10 mg/d, titrating to 40 mg/d vs placebo</td>
<td>132 boys</td>
<td>10–17</td>
<td>48 wk</td>
</tr>
<tr>
<td>Wiegman et al186 (2004)</td>
<td>Pravastatin 40 mg/d vs placebo</td>
<td>214</td>
<td>8–18</td>
<td>2 y</td>
</tr>
<tr>
<td>Bile-acid–binding resins</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tonstad et al187 (1996)</td>
<td>Colestipol 10 g/d or 5 g twice daily vs placebo</td>
<td>66 adolescents</td>
<td>NR</td>
<td>8 wk</td>
</tr>
<tr>
<td>Tonstad et al188 (1996)</td>
<td>Cholestyramine titrating up from 4 to 8 g/d vs placebo</td>
<td>96 boys</td>
<td>6–11</td>
<td>1 y</td>
</tr>
</tbody>
</table>

NR indicates not reported; TG, triglycerides; ↑, significant increase; ↓, significant decrease; ○, no significant change.
<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>Intervention(s)</th>
<th>Population</th>
<th>Duration of Trial</th>
<th>Significant Changes vs Control</th>
<th>Quality Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diet</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DISC Collaborative Research Group (1995)</td>
<td>Family-oriented behavioral intervention to promote dietary adherence vs usual care</td>
<td>663 with elevated LDL</td>
<td>4-10 y</td>
<td>↓</td>
<td>Good</td>
</tr>
<tr>
<td>Gold et al. (1991)</td>
<td>Oat bran–supplemented cereal within AHA Step 1 diet vs cereal with no oat bran</td>
<td>49 with TC &gt;185 mg/dL</td>
<td>10 y (mean)</td>
<td>↓ (year 1 only)</td>
<td>Poor</td>
</tr>
<tr>
<td>Kuehle et al. (1993)</td>
<td>Four 90-min family-oriented nutrition sessions vs one 90-min session</td>
<td>295 with TC &gt;185</td>
<td>2-15 y</td>
<td>↓</td>
<td>Poor</td>
</tr>
<tr>
<td>Obarsanek et al. (2001)</td>
<td>Counseling intervention (same as DISC above) vs usual care</td>
<td>663 with TC &gt;185</td>
<td>8-10 y</td>
<td>↓</td>
<td>Good</td>
</tr>
<tr>
<td>Shannon et al. (1994)</td>
<td>PCAT: 10 talking-book lessons and follow-up paper and pencil games for children with a manual for parents vs 45- to 60-min counseling session with parent, child, and registered dietitian and take-home print materials for both</td>
<td>261 with elevated LDL</td>
<td>4-10 y</td>
<td>↓</td>
<td>Good</td>
</tr>
<tr>
<td>Stallings et al. (1993)</td>
<td>PCAT: 10 sessions total, 1 per week completed in home by child and parents vs usual care</td>
<td>44 with LDL 90th–99th percentile</td>
<td>4-10 y</td>
<td>↓</td>
<td>Poor</td>
</tr>
<tr>
<td>Williams et al. (1995)</td>
<td>Fiber cereal with 3.2 g of soluble fiber per serving (dose = 1 box of cereal per d for 3 wk, then 2 boxes per d), with children aged 2–5 y consuming only 1 box per d throughout study, compared to placebo cereal with 0.5 g of fiber</td>
<td>58 with TC &gt;170 mg/dL and LDL &gt;110 mg/dL</td>
<td>2-11 y</td>
<td>↓</td>
<td>Poor</td>
</tr>
<tr>
<td><strong>Exercise</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boreham et al. (2000)</td>
<td>7-wk stair-climbing program vs no change in activity</td>
<td>25 sedentary females</td>
<td>18-22 y</td>
<td>↑</td>
<td>Fair</td>
</tr>
<tr>
<td>Ferguson et al. (1999)</td>
<td>Exercise program 5 d/wk, 40 min/d (children were paid $1 per session and given prizes for maintaining a heart rate &gt;150 beats per min) vs no exercise program</td>
<td>81 obese children</td>
<td>9.5 y (mean)</td>
<td>↑</td>
<td>Fair</td>
</tr>
<tr>
<td>Kang et al. (2002)</td>
<td>Physical activity training with lifestyle intervention 5 d/wk vs lifestyle intervention alone</td>
<td>80 obese children</td>
<td>13-16 y</td>
<td>↑</td>
<td>Poor</td>
</tr>
<tr>
<td>Linder et al. (1983)</td>
<td>Physical conditioning program vs usual activities</td>
<td>50 healthy boys</td>
<td>11-17 y</td>
<td>↓</td>
<td>Fair</td>
</tr>
<tr>
<td>Savage et al. (1986)</td>
<td>Walking/jogging/running 3 d/wk (1.6 km per session) high intensity (heart rate = 75% of V̇O₂max) vs low intensity (heart rate = 40% of V̇O₂max)</td>
<td>663 boys</td>
<td>8-9 y (mean)</td>
<td>↑</td>
<td>Fair</td>
</tr>
<tr>
<td>Stergioulas et al. (1998)</td>
<td>Four 60-min sessions per wk vs no specific training program</td>
<td>58 sedentary boys</td>
<td>10-14 y</td>
<td>↑</td>
<td>Fair</td>
</tr>
<tr>
<td><strong>Diet and Exercise</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Becque et al. (1988)</td>
<td>(1) Diet and behavior change: met with dietitian and behavior therapist 1 d/wk; (2) exercise plus diet and behavior change: same as above, with exercise program 50 min for 3 d/wk; and (3) no change in activity or diet</td>
<td>36 overweight adolescents</td>
<td>13 y (mean)</td>
<td>↑</td>
<td>Fair</td>
</tr>
<tr>
<td>Epstein et al. (1989)</td>
<td>Diet of 3800–5000 kJ/d monitored by a nutritionist; information on diet, exercise, stimulus control, reinforcement, modeling, and contingencies contract presented to parents and their children in 8 weekly sessions followed by 4 monthly sessions</td>
<td>56 obese (&gt;20% of ideal weight) children</td>
<td>8-12 y</td>
<td>↑</td>
<td>Poor</td>
</tr>
<tr>
<td>Walter et al. (1985)</td>
<td>&quot;Know Your Body&quot; curriculum yearly, taught 2 h/wk by usual classroom teacher, vs standard curriculum</td>
<td>1115</td>
<td>4th-graders</td>
<td>↑</td>
<td>Fair</td>
</tr>
</tbody>
</table>

↑ indicates significant increase; ↓, significant decrease; O, no significant change; DISC, Dietary Intervention Study in Children; PCAT, Parent-Child AutoTutorial Program; NR, not reported; TG, triglycerides.

This trial reported significant pre-experimental differences between groups in HDL levels (P < .05).
**Exercise Treatment**

Six studies evaluated exercise in normal-weight or obese children with elevated lipid levels (Table 3). Three studies were limited by differential or low completion rates, small numbers of participants, or other deficiencies (lack of blinding, lack of intention-to-treat analysis). Four trials that compared supervised, scheduled sessions of aerobic and fitness training to control groups showed minimal or no change in lipid levels compared with control groups. Two trials showed improvements in HDL-C for the exercise-intervention group compared with controls.

**Combination Diet and Exercise Treatment**

Three trials evaluated combined regimens of diet and exercise (Table 3). Although all the interventions showed some improvement in lipid levels, a group that undertook exercise, diet, and behavior changes had a 23% increase in HDL-C levels compared with both the diet-plus-behavior-change group and the control group.

**Key Question 9: What Are the Adverse Effects of Drug, Diet, Exercise, and Combination Therapy in Children/Adolescents?**

**Drug Treatment**

Information about adverse events was reported in 15 studies of statins, 22 studies of bile-acid–binding resins, and 8 studies of various other drugs or drug combinations. Studies used RCT, open-label–trial, and observational designs.

- Statins were associated with increased alanine aminotransaminase and/or aspartate aminotransferase levels in some, but not all, studies.
- Reports of elevated creatine kinase levels were similarly conflicting.

- Bile-acid–binding resins were associated with gastrointestinal complaints (8%–26%), such as flatulence and constipation, and unpalatability (up to 50%).

One study of cholestyramine reported transient increases in lactate dehydrogenase and abnormalities in aspartate aminotransferase levels that persisted for 6 months, but others showed normal liver-function test results. Growth was reported to be normal in 9 studies.

- One study reported a child whose height for age dropped below −2 SD while on colestipol (1 SD = 2.4 cm), whereas growth was normal in all other children in the study. Sexual maturation was followed over 4.3 years of treatment and found to be normal.

- Two studies of niacin reported increased liver enzyme levels (6 of 21 children in 1 study) and multiple other symptoms such as flushing, abdominal pain, nausea, and headache. There are also case reports of hepatitis and hepatotoxicity with the use of niacin.

**Low-Fat Diet**

Nineteen studies of dietary fat restriction reported effects on growth, nutrient intake, laboratory safety parameters, or other adverse effects.

Twelve studies reported normal height growth, although weight loss occurred among 3 children in 2 of these studies. In 1 study, growth failure occurred in 8 (20%) of 40 children with dyslipidemia, 3 (7.5%) of whom had nutritional dwarfing and no progression of puberty. In this study, families were unsupervised in the implementation of low-fat, low-cholesterol diets for a period up to 4.5 years; those with nutritional dwarfing had longer periods of time between diagnosis and formal dietary assessment and counseling. Failure to thrive has been demonstrated in children under 2 years of age who eat fat-restricted diets; these diets are not recommended for children in this age group.

**Dietary Supplements**

Fourteen studies provided information about adverse effects of various dietary supplements. Two children (4% of the treatment group) reported abdominal discomfort using fiber tablets (containing 50% wheat bran and 50% pectin) administered at 100 to 150 mg/kg per day. There were no adverse effects with psyllium fiber in 2 other studies. Other adverse effects of dietary supplements were mild or transient.

**Exercise**

A school-based program examined the effect of supervised exercise training on the lipid profiles of normal prepubertal children and reported 100% adherence and no adverse effects. In another study, treadmill tests elicited an exaggerated blood-pressure response in boys with dyslipidemia.

**Key Question 10: Does Improving Dyslipidemia in Childhood Reduce the Risk of Dyslipidemia in adulthood?**

No studies were identified that directly evaluated whether treatment of idiopathic dyslipidemia in childhood reduces risk of dyslipidemia in adulthood.

**CONCLUSIONS**

Although many studies have resolved the various aspects of dyslipidemia in children, few key questions about screening have been addressed (Table 5). Studies are not available that address the overarching key question about efficacy of screening children and adolescents for dyslipidemia in delaying the onset and reducing the incidence of CHD-related events (key question 1), effectiveness of treatments (drug, diet, exercise, and combination) on reducing incidence of adult dyslipidemia or delaying the onset and reducing the risk of CHD-related events (key question 4), or whether improving dyslipi-
<table>
<thead>
<tr>
<th>Author, year, title</th>
<th>Drug</th>
<th>Population</th>
<th>Duration of Trial</th>
<th>Adverse Effects of Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Statins</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McCrindleet al (2003)</td>
<td>Atorvastatin</td>
<td>187</td>
<td>10–17 y</td>
<td>26 wk</td>
</tr>
<tr>
<td>Clausset al (2005)</td>
<td>Lovastatin</td>
<td>54 girls</td>
<td>10–17 y</td>
<td>24 wk</td>
</tr>
<tr>
<td>Lambertet al (1996)</td>
<td>Lovastatin</td>
<td>69 boys</td>
<td>&lt;18 y</td>
<td>8 wk</td>
</tr>
<tr>
<td>Stein et al (1999)</td>
<td>Lovastatin</td>
<td>132</td>
<td>13 y (mean)</td>
<td>48 wk</td>
</tr>
<tr>
<td>Wiegmanet al (2004)</td>
<td>Pravastatin</td>
<td>214</td>
<td>8–18 y</td>
<td>2 y</td>
</tr>
<tr>
<td>Hedmanet al (2003)</td>
<td>Pravastatin</td>
<td>20</td>
<td>4–15 y</td>
<td>8 wk</td>
</tr>
<tr>
<td>Knipscheeret al (1996)</td>
<td>Pravastatin</td>
<td>72</td>
<td>12 y (mean)</td>
<td>12 wk</td>
</tr>
<tr>
<td>Coutureet al (1998)</td>
<td>Simvastatin</td>
<td>63</td>
<td>8–17 y</td>
<td>6 wk</td>
</tr>
<tr>
<td>de Jonghet al (2002)</td>
<td>Simvastatin</td>
<td>69</td>
<td>9–18 y</td>
<td>28 wk</td>
</tr>
<tr>
<td>Dirsamere t al (2003)</td>
<td>Simvastatin</td>
<td>20</td>
<td>10–17 y</td>
<td>18 mo</td>
</tr>
<tr>
<td>Stefanuttiet al (1999)</td>
<td>Simvastatin</td>
<td>16</td>
<td>7–12 y</td>
<td>12 mo</td>
</tr>
<tr>
<td>Various or unspecified statins</td>
<td>Various statins</td>
<td>22 professional athletes</td>
<td>15–27 y</td>
<td>8 y</td>
</tr>
<tr>
<td>Sinzinger and O’Grady (2004)</td>
<td>Various statins</td>
<td>69</td>
<td>10–18 y</td>
<td>NR</td>
</tr>
<tr>
<td>Curtis et al (1991)</td>
<td>Cholestyramine</td>
<td>1</td>
<td>7 y</td>
<td>2 y</td>
</tr>
<tr>
<td>Farah et al (1977) and Fashe et al (1977)</td>
<td>Cholestyramine</td>
<td>20</td>
<td>4–23 y</td>
<td>16 d</td>
</tr>
</tbody>
</table>

**Clinical Effects**

- None observed; no effect on sexual development.
- Abdominal pain (2), diarrhea (1), nausea (1), headache (1), and amenorrhea (1), all resolved with patient continuing medication.
- None observed.
- Abdominal pain (1), loose stools (1), headache (4), sleep disturbance (2), muscle tenderness or pain at rest (1), and muscle tenderness or pain associated with physical training (1).
- None observed.
- None observed.
- Rash, nose-bleeding, headache, nausea/vomiting, and abdominal pain.
- None observed.
- None observed.
- Transient headache (2), myalgia (1) for 2 wk, and transient gastrointestinal complaints (2).
- None observed.
- None observed.
- Muscle pain reported in 84% of periods of statin therapy (mean time of onset was 8.3 d).
- Loss of dental enamel noted (presumed caused by low pH 2.4 of cholestyramine mixed with Kool-Aid for administration).
- Febrile gastroenteritis (1) after 7-d treatment resulting in discontinuation of therapy.

**Laboratory Effects**

- Increased AST and ALT levels (1% of patients); no withdrew or stopped medication as a result of increased transaminase levels.
- Transient decreased HCT.
- None observed.
- Asymptomatic elevations in CK level (3).
- No effect on muscle or liver enzyme levels.
- No effect on serum ALT, CK, or creatinine levels.
- CK level abnormal in placebo (8), 5 mg/d (6), 10 mg/d (11), and 20 mg/d (3) groups; cortisol level abnormal in placebo (2), 5 mg/d (2), 10 mg/d (5), and 20 mg/d (3) groups.
- CK level abnormal in placebo (8), 5 mg/d (6), 10 mg/d (11), and 20 mg/d (3) groups; cortisol level abnormal in placebo (2), 5 mg/d (2), 10 mg/d (5), and 20 mg/d (3) groups.
- None observed.
- No significant effects on ALT, AST, and CK levels.
- Increased ALT (3), AST (3), and CK (1) levels.
- Slightly higher values of CK (2) transiently elevated ALT level and glucose challenge test (1).
- Transient increases in transaminase (1) and CK (2) levels.
- Elevated CK level in 3 subjects; no increase in liver enzyme levels.
- Serum calcium, phosphorus, folate, and B12 levels were within the reference ranges.
- Serum folate level decreased significantly in females; AST-level increases (2) persisted 6 mo; transient LDH increases (2); no fat-soluble vitamin malabsorption.
<table>
<thead>
<tr>
<th>Author, year, title</th>
<th>Drug</th>
<th>Population</th>
<th>Duration of Trial</th>
<th>Adverse Effects of Treatment</th>
<th>Clinical Effects</th>
<th>Laboratory Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glueck et al. (1973)</td>
<td>Cholestyramine</td>
<td>36</td>
<td>7–21</td>
<td>6 mo</td>
<td>None observed; normal growth</td>
<td>None observed</td>
</tr>
<tr>
<td>Glueck et al. (1974)</td>
<td>Cholestyramine</td>
<td>30 on diet + BABR</td>
<td>5–21 y</td>
<td>6-mo average follow-up</td>
<td>None observed; normal growth</td>
<td>Plasma vitamins A and E levels remained within the reference ranges</td>
</tr>
<tr>
<td>Glueck et al. (1977)</td>
<td>Cholestyramine</td>
<td>16</td>
<td>9–17 y</td>
<td>18 mo (16); 24 mo (12); 30–36 mo (7)</td>
<td>Persistent constipation (11); gritty sensation and poor palatability (5); chronic fatigue (1); dropouts after 2 y resulting from palatability</td>
<td>No effect on CBC, liver-function test results, or vitamin A and E, calcium, phosphorus, serum urea nitrogen, or fasting blood sugar levels</td>
</tr>
<tr>
<td>Glueck et al. (1986)</td>
<td>Cholestyramine</td>
<td>33</td>
<td>10.3 y (mean)</td>
<td>4.3 y</td>
<td>No effect on growth or sexual development; 1 competitive cross-country runner had persistently irregular periods</td>
<td>No effect on CBC, liver-function test results, or vitamin A and E, calcium, phosphorus, serum urea nitrogen, or fasting blood sugar levels</td>
</tr>
<tr>
<td>Koletzko et al. (1992)</td>
<td>Cholestyramine</td>
<td>35 on diet; 14 on diet + BABR</td>
<td>2–17 y</td>
<td>Mean: 17.5 mo (diet) and 27.9 mo (diet + BABR)</td>
<td>None observed; no effect on growth</td>
<td>None observed; normal growth</td>
</tr>
<tr>
<td>Liacouras et al. (1993)</td>
<td>Cholestyramine</td>
<td>87</td>
<td>106 y (mean)</td>
<td>Up to 62 mo</td>
<td>None observed; normal growth</td>
<td>None observed; normal growth</td>
</tr>
<tr>
<td>McCrindie et al. (1997)</td>
<td>Cholestyramine</td>
<td>40</td>
<td>10–18 y</td>
<td>28 wk</td>
<td>Persistent constipation (11); gritty sensation and poor palatability (5); chronic fatigue (1); dropouts after 2 y resulting from palatability</td>
<td>No effect on CBC, liver-function test results, or vitamin A and E, calcium, phosphorus, serum urea nitrogen, or fasting blood sugar levels</td>
</tr>
<tr>
<td>Tonstad et al. (1996)</td>
<td>Cholestyramine</td>
<td>96</td>
<td>6–11 y</td>
<td>1 y</td>
<td>No effect on growth; 1 case of intestinal obstruction caused by adhesions; unpalatability, headaches, and vomiting were reasons for withdrawals</td>
<td>No effect on growth; 1 case of intestinal obstruction caused by adhesions; unpalatability, headaches, and vomiting were reasons for withdrawals</td>
</tr>
<tr>
<td>Tonstad et al. (1998)</td>
<td>Cholestyramine</td>
<td>96</td>
<td>6–11 y</td>
<td>1 y</td>
<td>Unpalatability in both treatment and placebo groups</td>
<td>None observed; normal growth</td>
</tr>
<tr>
<td>West and Lloyd (1973)</td>
<td>Cholestyramine</td>
<td>19</td>
<td>1–14 y</td>
<td>Up to 20 mo</td>
<td>Unpalatability in both treatment and placebo groups</td>
<td>None observed; normal growth</td>
</tr>
<tr>
<td>West and Lloyd (1975)</td>
<td>Cholestyramine</td>
<td>18</td>
<td>1–14 y</td>
<td>1 to 2.5 y</td>
<td>No child developed diarrhea; no effect on growth</td>
<td>None observed; normal growth</td>
</tr>
<tr>
<td>West et al. (1975)</td>
<td>Cholestyramine</td>
<td>45</td>
<td>1–16 y</td>
<td>2–8 y</td>
<td>No child developed diarrhea; no effect on growth</td>
<td>None observed; normal growth</td>
</tr>
<tr>
<td>West et al. (1980)</td>
<td>Cholestyramine</td>
<td>35</td>
<td>1–17 y</td>
<td>1–8 y</td>
<td>Adherence was poor because of unpalatability</td>
<td>None observed; normal growth</td>
</tr>
<tr>
<td>West et al. (1983)</td>
<td>Colestipol</td>
<td>33</td>
<td>NR</td>
<td>16 wk</td>
<td>None observed; normal growth</td>
<td>None observed; normal growth</td>
</tr>
<tr>
<td>Hansen et al. (1992)</td>
<td>Colestipol</td>
<td>30</td>
<td>1–17 y</td>
<td>8.5 y (diet); 5.5 y (diet followed by diet + BABR)</td>
<td>Withdrawals because of unpalatability (5); 1 child’s height/age decreased below −2 SD; growth was normal in other children</td>
<td>None observed; normal growth</td>
</tr>
<tr>
<td>Harvengt and Desager (1976)</td>
<td>Colestipol</td>
<td>3</td>
<td>6–18 y</td>
<td>Up to 36 mo</td>
<td>Mild gastrointestinal complaint (flatulence, constipation) during first 3 mo but disappeared despite continued treatment</td>
<td>None observed; normal growth</td>
</tr>
</tbody>
</table>

NR: Not reported.
<table>
<thead>
<tr>
<th>Author, year, title</th>
<th>Drug</th>
<th>Population</th>
<th>Duration of Trial</th>
<th>Clinical Effects</th>
<th>Adverse Effects of Treatment</th>
<th>Laboratory Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>McCrindie et al. (2002)</td>
<td>Colestipol</td>
<td>N = 40, Age 9–18 y,</td>
<td>36 wk</td>
<td>Constipation (18%), stomachache (2%), headache (1%), and muscle aches (6%)</td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td>Schwarz et al. (1980)</td>
<td>Colestipol</td>
<td>N = 23, Age 5–17 y,</td>
<td>Up to 24 mo</td>
<td>Poor palatability (6), Raynaud disease phenomenon occurred during therapy (1), but treatment continued without recurrence</td>
<td></td>
<td>Serum vitamins A and E levels decreased significantly after 18–24 mo of colestipol</td>
</tr>
<tr>
<td>Tonstad et al. (1996)</td>
<td>Colestipol</td>
<td>N = 66, Age 13.2 y (mean),</td>
<td>52 wk</td>
<td>Gastrointestinal adverse effects (8) including constipation, dyspepsia, flatulence, nausea, decreased appetite, and abdominal pain; growth was normal</td>
<td></td>
<td>Reduced serum folate level after 8 wk; decreased serum vitamin E and carotenoid levels; decreased vitamin D levels (not significant) in subjects who were more compliant after 1 y</td>
</tr>
<tr>
<td>Tonstad and Ose (1996)</td>
<td>Colestipol</td>
<td>N = 27, Age 10–16 y,</td>
<td>6 mo (colestipol); 6 y (mean, for diet)</td>
<td>No effect on growth; difficulty swallowing the tablets (2), flatulence (1), and abdominal discomfort (1)</td>
<td></td>
<td>NR</td>
</tr>
<tr>
<td>Other drugs and combinations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baker et al. (1982)</td>
<td>Probucol</td>
<td>N = 7, Age 6–21 y,</td>
<td>15–21 mo</td>
<td>Nausea in 1 patient; no effect on growth and development</td>
<td></td>
<td>None observed</td>
</tr>
<tr>
<td>Becker et al. (1992)</td>
<td>Sitosterol and bezafibrate, in sequence and in combination</td>
<td>N = 7, Age 8.4 y (mean),</td>
<td>3 mo sitosterol; 3 mo bezafibrate; 24 mo sitosterol + bezafibrate</td>
<td>Decreased appetite for the first 2 wk on sitosterol (2)</td>
<td></td>
<td>Sitosterol: slight, significant decrease in hemoglobin (−5%) and ALP (−19%) levels; bezafibrate: ALP level remained lower, and iron increased by 26%; combination: transferrin increased 20% and reached abnormal levels in 2, and all other laboratory values were within reference ranges</td>
</tr>
<tr>
<td>Colletti et al. (1993)</td>
<td>Niacin</td>
<td>N = 21, Age 4–14 y,</td>
<td>1–19 mo, average 8.1 mo</td>
<td>18 of 21 patients reported some adverse effect: flushing (71%), itching (19%), abdominal pain (14%), nausea (14%), headache (14%), constipation (5%), and hepatitis (1)</td>
<td></td>
<td>Dose-related, reversible serum aminotransferase level elevations (6 [4 with crystalline and 2 with sustained-release form of niacin])</td>
</tr>
<tr>
<td>Malloy et al. (1978)</td>
<td>p-aminosalicylic acid</td>
<td>N = 20, Age 5–21 y,</td>
<td>6 mo</td>
<td>Mild gastric irritation that remitted with oral antacid treatment</td>
<td></td>
<td>Normal AST, ALT, ALP, bilirubin, and glucose levels in fasting serum; thyrotropin and thyroxine levels within the reference ranges</td>
</tr>
<tr>
<td>McDuffie et al. (2002)</td>
<td>Orlistat</td>
<td>N = 20, Age 146 y (mean),</td>
<td>3 mo</td>
<td>Gastrointestinal effects related to increased fat excretion that resolved within the first 6 wk of treatment; 1 subject withdrew because of intolerance of adverse effects</td>
<td></td>
<td>Decreased 25-hydroxy vitamin D levels at 1 mo; 3 subjects required additional vitamin D supplementation despite the prescription of a daily multivitamin containing vitamin D</td>
</tr>
<tr>
<td>Stein (1989)</td>
<td>Diet + drug or combined drugs: BABR, BABR + niacin, lovastatin or simvastatin</td>
<td>N = 30, Age 1–20 y,</td>
<td>1–9 y</td>
<td>None observed</td>
<td></td>
<td>Resin + niacin together produced elevated AST and ALT levels, decreased albumin levels, and clinical symptoms of hepatotoxicity (1)</td>
</tr>
<tr>
<td>Steinmetz et al. (1981)</td>
<td>Fenofibrate</td>
<td>N = 17, Age 4–19 y,</td>
<td>18 mo</td>
<td>None observed</td>
<td></td>
<td>Increased AL and AST (4) levels; decreased uric acid, bilirubin, inorganic phosphates, ALP, and GGT levels</td>
</tr>
<tr>
<td>Wheeler et al. (1985)</td>
<td>Bezafibrate</td>
<td>N = 14, Age 4–15 y,</td>
<td>3 mo</td>
<td>None observed; no effect on growth; all subjects declared preference for this drug over cholestyramine</td>
<td></td>
<td>Increased alkaline phosphatase level (1) and transient rise in ALT level (1)</td>
</tr>
</tbody>
</table>

(*) indicates the number of participants who experienced the effect; ALP, alkaline phosphate; ALT, alanine aminotransaminase; AST, aspartate aminotransferase; BABR, bile-acid–binding resin; CBC, complete blood count; CK, creatine kinase; DHEAS, dehydroepiandrosterones; GGT, γ-glutamyl transpeptidase; HCT, hematocrit; LDH, lactate dehydrogenase; NR, not reported.
<table>
<thead>
<tr>
<th>Arrow</th>
<th>Key Question</th>
<th>Quality of Evidence</th>
<th>Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Is screening for dyslipidemia in children effective in delaying the onset and</td>
<td>No evidence</td>
<td>No evidence</td>
</tr>
<tr>
<td></td>
<td>reducing the incidence of CHD-related events?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>What is the accuracy of screening for dyslipidemia in identifying children/</td>
<td>See subquestions</td>
<td>See subquestions below</td>
</tr>
<tr>
<td></td>
<td>adolescents at increased risk of CHD-related events?</td>
<td>below</td>
<td></td>
</tr>
<tr>
<td>2a</td>
<td>What are abnormal lipid values in children?</td>
<td>Fair to Poor</td>
<td>Normal values for lipids in children are currently defined according to population levels</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(percentiles). NCEP recommendations are based on LRC data, which defines the 95th percentile</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>for TC as 200 mg/dL and for LDL as 130 mg/dL. There are more recent studies suggesting that</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>age, gender, racial differences, and temporal trends shift these cut points. The NCEP has</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>defined levels of LDL for which drug treatment (LDL \geq 190 mg/dL or LDL \geq 160 mg/dL with</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>family history of early CHD), additional evaluation, diet therapy and testing (LDL \geq 130</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>mg/dL), and diet therapy with increased surveillance (LDL = 110–129 mg/dL) are recommended.</td>
</tr>
<tr>
<td>2b</td>
<td>What are appropriate tests? How well do screening tests (nonfasting TC, fasting</td>
<td>Poor</td>
<td>The most appropriate test is one that accurately predicts future risk and benefit from</td>
</tr>
<tr>
<td></td>
<td>TC, fasting lipid analysis) identify children/adolescents with dyslipidemia?</td>
<td></td>
<td>treatment. In the general population of children there have not been adequate studies to</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>determine these characteristics. Data from few studies suggest that TC &gt;95th percentile predicts</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LDL &gt;95th percentile with 44%–69% sensitivity. TC minus HDL might be a more sensitive test but</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>has not been extensively evaluated. A single TC measurement is inadequate to classify children</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>and adolescents into NCEP risk categories with 95% confidence.</td>
</tr>
<tr>
<td>2c</td>
<td>How well do lipid levels track from childhood to adulthood?</td>
<td>Good</td>
<td>Serial correlations measured in individual children over time are higher for TC (r = 0.38–0.78)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>and LDL r = 0.4–0.7) than for HDL and TG. Approximately 40%–55% of children with elevated</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>lipid levels (by percentile) will continue to have elevated lipid levels on follow-up.</td>
</tr>
<tr>
<td>2d</td>
<td>What is the accuracy of family history in determining risk?</td>
<td>Good</td>
<td>Multiple good-quality studies evaluating the use of family history as a diagnostic test for</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>dyslipidemia in children using varied and large populations demonstrate that family history is</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>an imperfect screening tool for detecting dyslipidemia among children.</td>
</tr>
<tr>
<td>2e</td>
<td>What are other important risk factors?</td>
<td>Good for family</td>
<td>Evidence from epidemiologic cross-sectional and cohort studies establishes statistical</td>
</tr>
<tr>
<td></td>
<td></td>
<td>history; good for</td>
<td>associations between elevations in lipid levels and family history and overweight. There is</td>
</tr>
<tr>
<td></td>
<td></td>
<td>obesity; poor for</td>
<td>inadequate evidence to show the magnitude of the effect of overweight on lipid levels or the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>all other risk</td>
<td>impact that incorporating weight measures into a screening tool could have. Multiple other risk</td>
</tr>
<tr>
<td></td>
<td></td>
<td>factors.</td>
<td>factors (diet, physical activity, aerobic capacity/fitness, puberty level, and smoking) have</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>not been evaluated adequately to assess their contribution to dyslipidemia in children or their</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>usefulness as screening tools.</td>
</tr>
<tr>
<td>2f</td>
<td>What are effective screening strategies for children/adolescents (including</td>
<td>Poor</td>
<td>Currently recommended screening strategies have limited diagnostic accuracy, low adherence</td>
</tr>
<tr>
<td></td>
<td>frequency of testing, optimal age for testing)?</td>
<td></td>
<td>to guidelines by providers, and limited compliance by parents and children. No trials compare</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>strategies of screening in children. No studies address the frequency and optimal age for</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>testing. Studies demonstrate lack of parental compliance with screening and follow-up</td>
</tr>
<tr>
<td>3</td>
<td>What are the adverse effects of screening including false-positive and false-</td>
<td>Fair</td>
<td>recommendations. Reasons for noncompliance include concern about test accuracy, lack of proof</td>
</tr>
<tr>
<td></td>
<td>negative results, labeling, etc?</td>
<td></td>
<td>that intervention makes a difference in children, concern about upsetting the child, refusal by</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>the child, inconvenience, or parental decision to institute a diet themselves and have child</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rechecked subsequently.</td>
</tr>
<tr>
<td>Arrow</td>
<td>Key Question</td>
<td>Quality of Evidence</td>
<td>Conclusions</td>
</tr>
<tr>
<td>-------</td>
<td>------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>4</td>
<td>In children and adolescents, what is the effectiveness of drug, diet, exercise, and combination therapy in reducing the incidence of adult dyslipidemia, and delaying the onset and reducing the incidence of CHD-related events and other outcomes (including optimal age for initiation of treatment)?</td>
<td>No evidence</td>
<td>No evidence</td>
</tr>
<tr>
<td>5–8</td>
<td>What is the effectiveness of drug, diet, exercise, and combination therapy for treating dyslipidemia in children/adolescents (including the incremental benefit of treating dyslipidemia in childhood)?</td>
<td>Good-quality studies with fair external validity for drug therapy; fair to poor for diet and exercise treatments</td>
<td>Statins are effective for reducing TC and LDL levels in children with FH. It is not clear how this efficacy translates to children with milder and/or nonfamilial forms of dyslipidemia. Diet supplements (psyllium, oat, sterol margarine) and counseling were marginally effective both in children and adolescents with FH/FCH and those without identified monogenic dyslipidemia. Exercise treatments showed minimal to no improvements in children without monogenic dyslipidemia.</td>
</tr>
<tr>
<td>9</td>
<td>What are the adverse effects of drug, diet, exercise, and combination therapy in children/adolescents?</td>
<td>Fair</td>
<td>Controlled and noncontrolled studies of treatment reported adverse effects of drug, diet, exercise, and combination therapy in children and adolescents. Statin drugs were associated primarily with elevations in liver enzymes (aspartate aminotransferase, alanine aminotransferase) and CK levels. bile-acid–binding resins were associated with adverse gastrointestinal effects and decreased levels of serum vitamins and minerals. Low-fat diets have been associated with growth retardation and nutritional dwarfing in 3 children who were placed on low-fat diets without formal advice and monitoring. Most studies show normal growth and development in children &gt;2 y old on monitored low-fat diets. Few adverse effects other than elevated blood pressure were noted with exercise. The duration of follow-up in these studies ranged from 10 d to 8 y. Studies were generally not of sufficient duration to determine long-term effects of either short or extended use.</td>
</tr>
<tr>
<td>10</td>
<td>Does improving dyslipidemia in childhood reduce the risk of dyslipidemia in adulthood?</td>
<td>No evidence</td>
<td>No evidence</td>
</tr>
</tbody>
</table>
demia in children and adolescents reduces the risk of adult dyslipidemia (key question 10).

Studies that evaluated risk factors are also limited. Risk factors that might contribute to a risk-assessment tool have not been studied adequately. Family-history questions are not standardized and have limited diagnostic accuracy. Evidence for risk factors other than family history for predicting dyslipidemia in children is strongest for overweight, but the magnitude of the effect of overweight on lipid levels, and the potential impact of incorporating overweight into a screening strategy for dyslipidemia, has not been addressed. Multiple other risk factors such as diet, physical inactivity, and aerobic capacity/fitness have not been evaluated adequately to assess their contribution to dyslipidemia or usefulness as screening tools either alone or in combination.

Currently recommended screening strategies have low adherence by providers and limited compliance by parents and children. No trials compared strategies by location, venue, age, or provider. No studies addressed the frequency and optimal age for testing. Adverse effects of screening for dyslipidemia have not been studied adequately.

Drug treatments for dyslipidemia in children have been studied only in children with FH or FCH, the population for whom these drugs are Food and Drug Administration-approved and recommended by the NCEP. Statins are effective for reducing TC and LDL-C levels in children with FH; it is not clear how this efficacy translates to children with milder and/or nonmonogenic dyslipidemia, and it is not known how frequently these medications are used in children without FH in practice. There are no trials with long-term follow-up for adult lipid outcomes or CHD-related events. Adverse effects of treatment are reported in controlled and noncontrolled studies of drug, diet, exercise, and combination therapy in children and adolescents. Studies were generally not of sufficient duration to determine long-term effects of either short or extended use.

Directions for future research should include identification of the impact of risk factors other than family history, such as overweight and physical inactivity, on lipid levels to develop risk-assessment strategies. Such tools may provide a better indication of actual risk and could facilitate screening by narrowing the number of children who require serum lipid testing. New vascular markers such as apolipoprotein B and apolipoprotein A-I may prove to be useful for screening in children. Additional evaluation of arterial IMT as a risk factor identifiable in children and its usefulness as a screening tool may be warranted.

Randomized, controlled clinical trials of screening strategies to determine which are more effective than current practice in terms of both parental compliance and provider adherence to guidelines are important. Screening strategies for ensuring adequate assessment of minorities and those with unknown family history deserve attention. Continued follow-up of currently established cohorts to assess the impact of screening for dyslipidemia in childhood on adult CHD outcomes is important.

More rigorous study designs, enrollment of larger population-based samples, and systematic reporting of adverse effects could improve studies of dyslipidemia treatments. Long-term follow-up of children treated with statins to determine the impact of sustained improvement of lipid levels in childhood on adult lipid levels, adult CHD outcomes, and long-term safety will help further assess the efficacy and safety of treatment options. The effect of exercise on lipid levels should be evaluated further, particularly in children with lipid levels >95th percentile. Standardized methods for collecting and reporting adverse effects in treatment trials would facilitate combining data across trials and lead to a more thorough understanding of the risks of treatment among children and adolescents.

ACKNOWLEDGMENTS

This study was conducted by the Oregon Evidence-based Practice Center under contract to the Agency for Healthcare Research and Quality (contract 290-02-0024, task order 2).

Agency staff, USPSTF members, and content experts reviewed interim reports. We thank Andrew Hamilton, MLS, MS, for conducting the literature searches; expert reviewers for commenting on draft versions of the systematic evidence synthesis; and Agency for Healthcare Research and Quality Medical Officer Janelle Guirguis-Blake and members of the USPSTF who served as leads for this project, including Leon Gordis, MD, DrPH, Carol Loveland-Cherry, PhD, RN, FAAN, Albert Siu, MD, MSPH, and Barbara Yawn, MD, MSc, FAAFP.

REFERENCES


75. Lannon CM, Earp J. Parents’ behavior and attitudes toward


82. Davis PH, Dawson JD, Riley WA, Lauer RM. Carotid intimal-medial thickness is related to cardiovascular risk factors measured from childhood through middle age: the Muscatine Study. Circulation. 2001;104:2815–2819


PEDIATRICS Volume 120, Number 1, July 2007 e209
128. Heyden S, Schneider KA, Roberts KF. Effectiveness of education-screening on cholesterol levels of students. Ann Nutr Metab. 1991;35:71–76
141. Bastida S, Perea S, Sanchez-Muniz FJ. Do neonates with high serum cholesterol levels have a different high density lipoprotein composition? Eur J Pediatr. 1998;157:66–70
146. Anderssen GE, Brokhattingen K, Lous P. Familial hyperbetalipoproteinaemia in 9 children diagnosed as the result of cord blood screening for hyperlipoproteinemia in 10 000 Danish newborns. Arch Dis Child. 1979;54:691–694


228. Collette RB, Neufeld EJ, Rolf NK, McAuliffe Tl, Baker AL,


APPENDIX

USPSTF Quality Rating Criteria

Diagnostic accuracy studies

Criteria
- Screening test relevant, available for primary care, adequately described
- Study uses a credible reference standard, performed regardless of test results
- Reference standard interpreted independently of screening test
- Handles indeterminate results in a reasonable manner
- Spectrum of patients included in study
- Sample size
- Administration of reliable screening test

Definition of ratings on basis of above-listed criteria
- Good: evaluates relevant available screening test; uses a credible reference standard; interprets reference standard independently of screening test; reliability of test assessed; has few or handles indeterminate results in a reasonable manner; includes large number (>100) of broad-spectrum patients with and without disease
- Fair: evaluates relevant available screening test; uses reasonable although not best standard; interprets reference standard independent of screening test; moderate sample size (50–100 subjects) and a "medium" spectrum of patients
- Poor: has important limitations such as uses inappropriate reference standard, screening test improperly administered, biased ascertainment of reference standard, and/or very small sample size of very narrow selected spectrum of patients

RCTs and cohort studies

Criteria
- Initial assembly of comparable group
  - RCTs: adequate randomization, including concealment and whether potential confounders were distributed equally among groups
  - Cohort studies: consideration of potential confounders with either restriction or measurement for adjustment in the analysis
- Consideration of inception cohorts
- Maintenance of comparable groups (includes attrition, crossovers, adherence, contamination)
- Important differential loss to follow-up or overall high loss to follow-up
- Measurements: equal, reliable, and valid (includes masking of outcome assessment)
- Clear definition of interventions
- Important outcomes considered
- Analysis: adjustment for potential confounders for cohort studies or intention-to-treat analysis for RCTs

Definition of ratings on basis of above-listed criteria
- Good: meets all criteria: comparable groups as assembled initially and maintained throughout the study (follow-up at least 80%); reliable and valid measurement instruments are used and applied equally to the groups; interventions are spelled out clearly; important outcomes are considered; and appropriate attention to confounders in analysis
- Fair: studies will be graded "fair" if any or all of the following problems occur, without the important limitations noted in the "poor" category below
  - Generally comparable groups are assembled initially, but some question remains whether some (although not major) differences occurred in follow-up
  - Measurement instruments are acceptable (although not the best) and generally applied equally
  - Some, but not all, important outcomes are considered
  - Some, but not all, potential confounders are accounted for
- Poor: studies will be graded "poor" if any of the following major limitations exists
  - Groups assembled initially are not close to being comparable or maintained throughout the study
  - Unreliable or invalid measurement instruments are used or not applied at all equally among groups (including not masking the outcome assessment)
  - Key confounders are given little or no attention

Case-control studies

Criteria
- Accurate ascertainment of cases
- Nonbiased selection of cases/controls with exclusion criteria applied equally to both
- Response rate
- Diagnostic testing procedures applied equally to each group
- Measurement of exposure accurate and applied equally to each group
- Appropriate attention to potential confounding variable

Definition of ratings on basis of above-listed criteria
- Good: appropriate ascertainment of cases and nonbiased selection of case and control participants; exclusion criteria applied equally to cases and controls; response rate =80%; diagnostic procedures and measurements accurate and applied equally to cases and controls; and appropriate attention to confounding variables
- Fair: recent, relevant, without major apparent selection or diagnostic workup bias but with response rate <80% or attention to some, but not all, important confounding variables
- Poor: major selection or diagnostic workup biases, response rates <50%, or inattention to confounding variables

Screening for Lipid Disorders in Children: US Preventive Services Task Force Recommendation Statement

US Preventive Services Task Force

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

SUMMARY OF RECOMMENDATION
The US Preventive Services Task Force (USPSTF) concludes that the evidence is insufficient to recommend for or against routine screening for lipid disorders in infants, children, adolescents, or young adults (up to age 20) (I recommendation).

RATIONALE

Importance
There is good evidence that children with lipid disorders (dyslipidemia) are at risk for becoming adults with lipid disorders.

Detection
For children with familial dyslipidemia, the group most likely to benefit from screening, use of family history in screening may be inaccurate because of variability of definitions and unreliability of information. Serum lipid levels are accurate screening tests for childhood dyslipidemia, although many children with multifactorial types of dyslipidemia would have normal lipid levels in adulthood. Fifty percent of children and adolescents with dyslipidemia will have dyslipidemia as adults.

Benefits of Detection and Early Treatment*
Trials of statin drugs in children with monogenic dyslipidemia (defined below in “Clinical Considerations”) indicate improved total cholesterol (TC), low-density lipoprotein cholesterol (LDL-C), and high-density lipoprotein cholesterol (HDL-C) measures. For children with multifactorial types of dyslipidemia, there is no evidence that diet or exercise interventions in childhood lead to improved lipid profiles or better health outcomes in adulthood.

Harms of Detection and Early Treatment
Potential harms of screening may include labeling of children whose dyslipidemia would not persist into adulthood or cause health problems, although evidence is lacking. Adverse effects from lipid-lowering medications and low-fat diets, including potential long-term harms, have been inadequately evaluated in children.

USPSTF Assessment
The USPSTF was unable to determine the balance between potential benefits and harms for routinely screening children and adolescents for dyslipidemia.

*Critical evidence gap.
CLINICAL CONSIDERATIONS

- Dyslipidemias are abnormalities of lipoprotein metabolism and include elevations in TC, LDL-C, or triglyceride levels or deficiencies of HDL-C. These disorders can be acquired or familial; monogenic dyslipidemias are related to genetic conditions such as familial hypercholesterolemia in some individuals. Multifactorial dyslipidemias are caused by risk factors including environmental factors (obesity, diet) or currently unidentified genetic factors. This recommendation applies to all asymptomatic individuals from birth to age 20.

- Because abnormal lipid levels have been strongly associated with the risk of coronary heart disease (CHD) events in adulthood, and early identification and lipid-lowering intervention in certain populations of adults can prevent CHD events, much attention has been directed at screening individuals for dyslipidemia at young ages (eg, childhood). Among children and adolescents, 3 groups may be identified through screening: (1) children with undiagnosed monogenic dyslipidemias such as familial hypercholesterolemia; (2) those with undiagnosed secondary causes of dyslipidemia; and (3) those with multifactorial dyslipidemia (polygenetic or related to risk factors). However, the clinical health benefits shown in adults identified and treated for dyslipidemia have not been studied in children, which makes the role of screening children uncertain.

- Children and adolescents with diabetes may be at especially high risk for dyslipidemia and cardiovascular events. Screening children and adolescents with diabetes for dyslipidemia has been recommended by other groups as a part of appropriate care for these children.

- The use of family history as a screening tool for dyslipidemia has variable accuracy largely because definitions of a positive family history and lipid threshold values vary substantially. Screening using family history as defined by the National Cholesterol Education Program (NCEP) and the American Academy of Pediatrics has been shown to produce high rates of false-negative results.

- If clinicians choose to screen for dyslipidemia, the preferred screening tests are TC and HDL-C on nonfasting or fasting samples; calculating LDL-C requires fasting samples.

OTHER CONSIDERATIONS

- The effectiveness of treatment interventions (diet, exercise, lipid-lowering agents) in children with dyslipidemia (including multifactorial dyslipidemia) in improving health outcomes remains a critical research gap. Population-based screening studies or randomized, controlled trials (RCTs) following children and adolescents into adulthood after treatment interventions will be necessary to assess universal lipid screening in childhood or adolescence.

- Rising rates of childhood overweight may lead to a higher prevalence of dyslipidemia in childhood and adulthood. Continued tracking of dyslipidemia in all age groups will be important as the epidemiology of obesity evolves.

DISCUSSION

Epidemiology

Dyslipidemias are disorders of lipoprotein metabolism and include elevations in TC, LDL-C, or triglyceride levels or deficiencies of HDL-C. TC levels increase from birth, stabilize at ~2 years of age, peak before puberty, and then decline slightly during adolescence. Normal values for lipids in children and adolescents are currently defined according to population distributions of lipid levels from the Lipid Research Clinics Prevalence Study conducted in the 1970s. Dyslipidemia is commonly defined as TC > 200 mg/dL and LDL-C > 130 mg/dL; these values correspond to the 95th percentile observed in the Lipid Research Clinics study. More recent studies, including the National Health and Nutrition Examination Survey, indicate that age, sex, racial differences, and temporal trends shift these population-based cut points.

Although dyslipidemia in adults is an established risk factor for CHD on the basis of good-quality evidence from long-term prospective studies, the CHD risk attributable to dyslipidemia during childhood is unknown. Indirect evidence from the Bogalusa Heart Study, a long-term epidemiologic study of risk factors for CHD from birth through 31 years of age, showed a correlation between lipid levels and arterial fat deposition seen at autopsy; however, such evidence does not directly link childhood lipid levels to health outcomes. Epidemiologic studies in children establish a strong statistical association between childhood overweight and dyslipidemia. Other risk factors for dyslipidemia include an established family history for common familial dyslipidemias including familial hypercholesterolemia, familial combined hypercholesterolemia, familial defective apo-protein B, and familial hypertriglyceridemia. Secondary causes of dyslipidemia include diabetes, nephrotic syndrome, and hypothyroidism.

The USPSTF did not find direct evidence that screening for dyslipidemia leads to improvements in CHD-related mortality or overall mortality; therefore, it reviewed the evidence on accuracy of screening tests including family history, efficacy of treatment, and harms of screening and treatment in children.
Accuracy of Screening Tests

TC and HDL-C levels can be measured on nonfasting venous or capillary blood samples, LDL-C measurement requires fasting samples, and direct LDL-C can be measured on nonfasting venous samples. At least 2 measurements are necessary to ensure that true values are within 10% of the mean of the measurements. Fair-quality evidence shows that a value of TC minus HDL-C above the 95th percentile is 88% to 96% sensitive and 98% specific for detecting LDL-C above the 95th percentile is 88% to 96% sensitive and 98% specific for detecting LDL-C levels. Although use of family history presents a potential method to target serum lipid screening to a group of children and adolescents with higher risk for dyslipidemia, its use is limited. Family history is time-consuming to elicit accurately, it has been variably defined in the literature, and its use as a screening tool has been shown to miss substantial numbers (30%–60%, in general) of children with elevated lipid levels. Family-history definitions vary substantially among studies, as do lipid-detection thresholds; those studies that show higher sensitivities (~77%) have low specificities (≤55%).

Population-based estimates of the number of children who require serum lipid testing on the basis of positive family history may range from 25% to 55%, depending on definitions of family history and serum LDL cutoff values.

Accurate screening tests in children would be useful if childhood dyslipidemia correlated with adult CHD health outcomes or with adult dyslipidemia as an intermediate outcome and if treatment improved CHD outcomes. Serial correlations between lipid levels measured in individual children over time vary on the basis of the type of lipid level followed. On the basis of the evidence from 23 prospective cohort studies, correlations have been found to be higher for TC (r = 0.38–0.78) and LDL-C (r = 0.4–0.7) levels than for HDL-C (r = 0.0–0.8) and triglyceride (r = 0.1–0.58) levels, and good-quality evidence indicates that ~40% to 55% of children with elevated TC and LDL levels will continue to have elevated lipid levels on follow-up into adolescence and early adulthood. No studies examine tracking of lipid levels in those with risk factors for dyslipidemia (eg, childhood overweight).

Efficacy of Treatment

Treatment of childhood dyslipidemia has been shown to be effective in lowering lipid levels in select populations; however, no studies have addressed the effect of treatment on childhood or adult health outcomes (eg, CHD events). In those children with diagnosed monogenic dyslipidemia, a condition that has been associated with premature CHD events, no RCTs are likely to be completed to provide health outcomes in untreated controls. In this population of children with familial monogenic dyslipidemias (familial hypercholesterolemia or familial combined hyperlipidemia), good-quality evidence based on a meta-analysis of 9 RCTs demonstrated the effectiveness of statins in reducing intermediate outcomes: TC and LDL (percent mean reduction [95% confidence limits] from meta-analysis of trials: 24.4% [19.5, 29.2] for TC and 30.8% [24.1, 37.5] for LDL in 8 studies). Fair evidence based on 2 fair-quality trials shows that bile-acid–binding resins reduce lipid levels in children with monogenic dyslipidemia. RCTs of diet supplements (psyllium, oat, garlic extract, and sterol margarine) and advice show marginal improvements in lipid levels in children with monogenic dyslipidemia. There is fair-quality evidence that dietary counseling is associated with minimal improvements in lipid levels in children with monogenic and multifactorial dyslipidemias; however, these improvements may not be sustained after the counseling intervention ceases. There are no studies of physical activity interventions in those with monogenic dyslipidemia and fair-quality evidence in those with multifactorial dyslipidemia based on a meta-analysis of 6 trials that showed that physical activity interventions were associated with minimal to no improvement in lipid levels in children with multifactorial dyslipidemia (percent mean reduction [95% confidence limits] from meta-analysis of trials: 0% [−5.6, 5.6] for TC and 3.1% [−7.7, 1.5] for LDL-C reduction in 4 studies).

Harms of Screening and Treatment

There is poor-quality evidence on the adverse effects of screening. There are conflicting reports about behavioral difficulties in screened children and reports of parental noncompliance with recommendations for diet and follow-up. Studies have shown no increases in anxiety among screened children and adolescents. Fair-quality evidence on the harms of treatment is based on 81 controlled and noncontrolled studies of treatment that reported a variety of adverse effects of drug, diet, exercise, and combination therapy in children and adolescents. Lipid-lowering agents have been shown to cause elevations in creatine kinase and liver-function tests (statins), gastrointestinal adverse effects, and decreased absorption of vitamins and minerals (bile-acid resins). The adverse effects of long-term use of lipid-lowering agents (eg, for >20 years) have not been studied. There have been 3 reports of growth retardation and nutritional dwarfing in children on unmonitored diets; however, there are several reports of normal growth during monitored low-fat diet interventions. Physical activity interventions have had no reported harms in children without monogenic dyslipidemia, but an exaggerated blood pressure response was seen in children with monogenic dyslipidemias who were undergoing physical activity intervention.

RECOMMENDATIONS OF OTHERS

No professional organization recommends universal screening for dyslipidemia in children or adolescents. The NCEP report of the Expert Panel on Blood Choles-
terol Levels in Children and Adolescents recommends selective screening for children and adolescents with a family history of premature CHD or at least 1 parent with a high TC level (TC ≥ 240 mg/dL) in the context of regular health care. Optional cholesterol testing may be recommended in children and adolescents who are judged to be at higher risk independent of family history or parental hypercholesterolemia (eg, those who are overweight or have high-fat diets).

The American Academy of Pediatrics’ recommendations are based on this NCEP report and concur with its screening recommendations. The American College of Obstetricians and Gynecologists concurs with the NCEP recommendations for screening in adolescents. In 2003, the American Heart Association recommended performing targeted screening of fasting lipids in children >2 years of age with a family history of dyslipidemia or premature cardiovascular disease and in children for whom family history is unknown and other risk factors are present. In a 2007 update, the American Heart Association recommended, in addition, screening children who are overweight or obese.

APPENDIX 1: USPSTF RECOMMENDATIONS AND RATINGS

The USPSTF grades its recommendations according to 1 of 5 classifications (A, B, C, D, and I) reflecting the strength of evidence and magnitude of net benefit (benefits minus harms).

A. The USPSTF strongly recommends that clinicians provide [the service] to eligible patients. The USPSTF found good evidence that [the service] improves important health outcomes and concludes that benefits substantially outweigh harms.

B. The USPSTF recommends that clinicians provide [the service] to eligible patients. The USPSTF found at least fair evidence that [the service] improves important health outcomes and concludes that benefits outweigh harms.

C. The USPSTF makes no recommendation for or against routine provision of [the service]. The USPSTF found at least fair evidence that [the service] can improve health outcomes but concludes that the balance of benefits and harms is too close to justify a general recommendation.

D. The USPSTF recommends against routinely providing [the service] to asymptomatic patients. The USPSTF found at least fair evidence that [the service] is ineffective or that harms outweigh benefits.

I. The USPSTF concludes that the evidence is insufficient to recommend for or against routinely providing [the service]. Evidence that [the service] is effective is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

APPENDIX 2: USPSTF STRENGTH OF OVERALL EVIDENCE

The USPSTF grades the quality of the overall evidence for a service on a 3-point scale (good, fair, poor).

Good
Evidence includes consistent results from well-designed, well-conducted studies in representative populations that directly assess effects on health outcomes.

Fair
Evidence is sufficient to determine effects on health outcomes, but the strength of the evidence is limited by the number, quality, or consistency of the individual studies, generalizability to routine practice, or indirect nature of the evidence on health outcomes.

Poor
Evidence is insufficient to assess the effects on health outcomes because of limited number or power of studies, important flaws in their design or conduct, gaps in the chain of evidence, or lack of information on important health outcomes.

USPSTF MEMBERS

The following were the members of the USPSTF at the time this recommendation was finalized (for a list of current USPSTF members, go to www.ahrq.gov/clinic/uspsfab.htm): Ned Calonge, MD, MPH, chair, USPSTF (chief medical officer and state epidemiologist, Colorado Department of Public Health and Environment, Denver, CO); Diana B. Petitti, MD, MPH, vice-chair, USPSTF (senior scientific advisor for health policy and medicine, Kaiser Permanente Southern California, Pasadena, CA); Thomas G. DeWitt, MD (Carl Weihl professor of pediatrics and director of the Division of General and Community Pediatrics, Department of Pediatrics, Children’s Hospital Medical Center, Cincinnati, OH); Leon Gordis, MD, MPH, DrPH (professor, Epidemiology Department, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD); Kimberly D. Gregory, MD, MPH (director, Women’s Health Services Research and Maternal-Fetal Medicine, Department of Obstetrics and Gynecology, Cedars-Sinai Medical Center, Los Angeles, CA); Russell Harris, MD, MPH (professor of medicine, Sheps Center for Health Services Research, University of North Carolina School of Medicine, Chapel Hill, NC); Kenneth W. Kizer, MD, MPH (president and chief executive officer, National Quality Forum, Washington, DC); Michael L. LeFevre, MD, MSPH (professor, Department of Family and Community Medicine, University of Missouri School of Medicine, Columbia, MO); Carol Loveland-Cherry, PhD, RN (executive associate dean, Office of Academic Affairs, University of Michigan School of Nursing, Ann Arbor, MI); Lucy N. Marion, PhD, RN (dean and professor, School of Nurs-
ing, Medical College of Georgia, Augusta, GA); Virginia A. Moyer, MD, MPH (professor, Department of Pediatrics, University of Texas Health Science Center, Houston, TX); Judith K. Ockene, PhD (professor of medicine and chief of the Division of Preventive and Behavioral Medicine, University of Massachusetts Medical School, Worcester, MA); George F. Sawaya, MD (associate professor, Department of Obstetrics, Gynecology, and Reproductive Sciences and Department of Epidemiology and Biostatistics, University of California, San Francisco, CA); Albert L. Siu, MD, MSPH (professor and chairman, Brookdale Department of Geriatrics and Adult Development, Mount Sinai Medical Center, New York, NY); Steven M. Teutsch, MD, MPH* (executive director, Outcomes Research and Management, Merck & Company, Inc, West Point, PA); and Barbara P. Yawn, MD, MSc (director of research, Olmstead Research Center, Rochester, MN).

* Dr Teutsch was recused from the discussion and voting on this issue.

REFERENCES


Mollaret Meningitis Associated With an Intraspinal Epidermoid Cyst

Baoqin Gao, MD, Jun Yang, MD, Shougang Zhuang, MD, Yaxian Deng, MD, Weili Yang, MD, Yazhen Yu, MD, Yongjun Wang, MD, Lin Luo, MD, Ke Dai, MD

Departments of Pediatrics and Pediatric Neurology, Neurosurgery, and Neurology, Beijing Tiantan Hospital, Capital University of Medical Sciences, People’s Republic of China; Department of Medicine, Brown University School of Medicine, Providence, Rhode Island; Department of Neuropathology, Institute of Neuroscience, Chinese Academy of Sciences, People’s Republic of China

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

ABSTRACT

Mollaret meningitis, a benign recurrent aseptic disease, is known to be associated with intracranial epidermoid cysts. In this report, we describe a case of Mollaret meningitis caused by an intraspinal epidermoid cyst located at thoracic level 12. The patient’s clinical manifestations and cerebrospinal fluid features were similar to those with bacterial meningitis characterized by predominant polymorphonuclear leukocytes. However, Mollaret cells, not bacteria, were identified in the patient’s cerebrospinal fluid. The illness ceased after surgical resection of the cyst, and the cyst tissue was pathologically diagnosed as epidermoid. Therefore, an intraspinal epidermoid cyst can be etiologically associated with Mollaret meningitis and should be included in the differential diagnosis of recurrent aseptic meningitis.

MOLLARET MENINGITIS WAS initially described by Mollaret in 1944 as a form of aseptic meningitis without identifiable infecting agents. It is characterized as recurrent, aseptic, mild, and self-limiting. To date, ~50 cases have been reported in the world, and most have been associated with viral infection, especially herpes simplex type 2. Other reported etiologies include Vogt-Koyanagi syndrome, Harada syndrome, Behçet disease, allergic, systemic lupus erythematosus, familial Mediterranean fever, glioblastoma, and Whipple disease, as well as intracranial hydatid, sarcoidosis, and epidermoid cysts.

In 1962, Bruyn proposed criteria for the clinical diagnosis of Mollaret meningitis: (1) recurrent episodes of severe headache, meningismus, and fever; (2) cerebrospinal fluid (CSF) pleocytosis with large “endothelial” cells, neutrophils, and lymphocytes; (3) attacks separated by symptom-free periods that last weeks to months; (4) spontaneous remission of symptoms and signs; and (5) no causative etiologic agent. The so-called large endothelial cells were confirmed in subsequent studies to be enlarged monocytes/macrophages in origin. On the basis of these criteria, we recently identified a case of recurrent Mollaret meningitis that was associated with an intraspinal epidermoid cyst. Intraspinal epidermoid cysts in children frequently cause recurrent episodes of bacterial meningitis associated with a dermal sinus tract. The connection between intraspinal epidermoid cysts and the skin surface provides direct access for bacteria on the skin to reach the interior cysts and may result in meningitis. However, the epidermoid cyst in our patient was concealed in the spinal subarachnoid, and no such sinus tract was found. We believe this to be the first case of Mollaret meningitis secondary to an intraspinal epidermal cyst that was not associated with a communicating dermal sinus.

CASE REPORT

An 8-year-old girl presented with repeated episodes of malaise, fever, nausea, vomiting, and headache. In the previous 27 months, she had suffered 38 episodes of “meningitis,” each with similar clinical manifestations.

During the first episode, she had a temperature of 38.5 to 40.5°C, headache, and vomiting. Her physical examination revealed a stiff neck, positive Brudzinski sign, and positive Kernig sign. Her CSF was colorless with a hazy appearance and had a white blood cell

Key Words: Mollaret meningitis, epidermoid cyst

Abbreviations: CSF, cerebral spinal fluid; WBC, white blood cell; PMN, polymorphonuclear leukocyte; T12, thoracic level 12

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2053
doi:10.1542/peds.2006-2053

Accepted for publication Jan 25, 2007
Address correspondence to Baoqin Gao, MD, Tian Tan Xi Li 6#, Chong Wen District, Beijing 100050, People’s Republic of China. E-mail: gaobaoqin@sohu.com

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
(WBC) count of 1100/mm³. The WBC differential count was 12% lymphocytes, 2% monocytes, and 86% neutrophils. Glucose was 56 mg/dL, and protein was 106 mg/dL. A CSF stain was negative for acid-fast bacteria and fungus. A peripheral WBC count was 15 000/mm³, with 72% neutrophils and 18% lymphocytes. Blood and CSF bacterial cultures were negative. Brain MRI results were normal. The patient was admitted to a local hospital, and penicillin and cefazidime were given for presumed bacterial meningitis. The symptoms and signs disappeared without any local neurologic signs 7 days later. A repeat lumbar puncture performed on the 15th day of hospitalization revealed nearly normal CSF parameters (WBC count of 7/mm³, with 2 neutrophils). She was discharged 20 days postadmission. Subsequently, the girl experienced 37 episodes of “recurrent meningitis” with symptom-free intervals between episodes. Although she experienced these episodes of recurrent meningitis, the patient was otherwise healthy, and no neurologic sequelae were discovered. During every episode, the symptoms, signs, and recovery process were nearly the same. Laboratory results were similar to those of bacterial meningitis in many respects. There was a CSF pleocytosis with 1000 to 3500 WBC per mm³ and 56% to 99% neutrophils. The CSF protein was mildly elevated, and glucose was slightly decreased. The CSF usually became normal after 2 weeks of treatment. No known specific infectious cause was identified. Every episode was diagnosed as recurrent meningitis and treated with antibiotics, dexamethasone, and a dehydrating such as mannitol. The symptom-free periods ranged from 2 to 6 months in the early period. During the last 3 months of her illness, the attacks became more frequent, with intervals of 3 to 7 days between episodes.

On May 11, 2005, she came to our hospital complaining of headache, fever, and stiff neck. Physical examination on admission revealed a temperature of 39.6°C, heart rate of 112 beats per minute, respiratory rate of 25 breaths per minute, and blood pressure of 96/65 mm Hg. No rash, lymphadenopathy, or skin sinuses were found. Neurologic examination revealed a stiff neck, positive Brudzinski sign, and positive Kernig sign. No other focal neurologic signs were discovered. A lumbar puncture was performed on admission and revealed a hazy CSF that contained no red blood cells and 2520 WBCs per mm³, with 99% polymorphonuclear leukocytes (PMNs) and 1% monocytes. The glucose level was 41 mg/dL, chloride level was 121 mmol/L, and protein level was 103.3 mg/dL. WBCs and protein in CSF became normal after 3 weeks. CSF bacterial and viral cultures were all negative (data not shown). Neither cryptococcal antigen nor virus antibodies were detected (data not shown). A peripheral-blood WBC count was 4300/mm³, with 37% neutrophils, 54% lymphocytes, and 9% monocytes.

The cytologic study of the CSF revealed the presence of Mollaret-like cells. Phenotypically, Mollaret-like cells were larger than the adjacent PMNs and appeared with round, oval shapes or were kidney shaped with deep nuclear clefts, resulting in unusual nuclear shapes. Some of the Mollaret-like cells contained rich cytoplasm and vacuoles. The degenerated cells appeared as the so-called ghost cells in the slide background (Fig 1).

In addition, her blood immunoglobulin G level was 27.5 g/L (reference range: 9.4–16.19 g/L); tubercle bacillus polymerase chain reaction was negative; blood CD, CD8, and antistreptolysin O levels were within the reference ranges; and a purified protein derivative test was negative. Microbiologic and serologic testing results were all negative (data not shown).

MRI revealed a well-defined round shape of T2-weighted higher signal ventrally situated opposite to thoracic level 12 (T12). Neither vertebral defects nor gadolinium enhancement were found. In addition, the cyst did not present with spinal cord compression (Fig 2). Head MRI results were normal (data not shown).

The patient received antibiotics (ceftriaxone) and other treatments (ie, dexamethasone and mannitol). The patient’s temperature returned to a normal range, and other symptoms disappeared within 7 days; CSF became normal on day 15.

Subsequently, surgery was performed to remove the cyst. On exposure of the spinal cavity at T12, an intact, translucent, thin-walled cyst (10 × 8 × 5 mm) was observed (Fig 3A, arrow). The resected cyst had a white, firm wall without fistula through the outside or to other tissues or organs. The cyst contained 1 mL of ivory clouded liquid. Microscopically, the inner layer of the cyst wall consisted of squamous epithelium. The external layer was formed by connective tissue. No keratin or...
inflammation reaction was observed in the cyst wall (Fig 3 B and C). Cellular debris and keratin were observed within the cyst, but there were no inflammatory cells. Gram-staining and bacterial cultures of the cyst liquid were negative (data not shown), as were acid-fast bacteria and fungi cultures. The cyst tissue was pathologically diagnosed as epidermoid.

Postoperative MRI showed that the cyst was removed completely, and the spinal cord was integrated (Fig 2). The child recovered completely after the operation, and no additional episodes of meningitis have occurred during the next 15 months.

DISCUSSION
Mollaret meningitis was first described as a benign recurrent aseptic meningitis of suspected viral etiology. In particular, herpes simplex virus type 2 was frequently reported to be associated with this disease.2–5 However, our laboratory testing (blood and CSF examination) in this patient did not reveal the presence of herpes simplex virus type 2 or other viruses. Instead, MRI revealed a spinal cyst at T12, with pathologic epidermoid confirmation (Figs 2 and 3). Its association with Mollaret meningitis was demonstrated by the cessation of symptoms and signs after surgical removal of the cyst. Therefore, in suspected cases of Mollaret meningitis, imaging of the spinal cord should be considered despite the absence of spinal cord symptomatology.

Our patient had a similar clinical presentation and CSF profile to other cases reported to be caused by intracranial cysts. Analysis of the CSF in this patient and other patients with Mollaret meningitis associated with epidermoid cysts revealed a feature distinct from viral Mollaret during the acute phase (Tables 1 and 2). Although the CSF of patients with Mollaret meningitis caused by herpes simplex virus type 2 displays the same features as other viral meningitis (a mild increase in WBCs and neutrophils), the CSF of those with epidermoid cysts shows more marked CSF pleocytosis with a predominance of PMNs. In addition, frequent recurrence of meningitis without neurologic sequela is another clinical feature of Mollaret meningitis that is different from bacterial meningitis. Thus, we suggest that, among patients with clinical symptoms indicative of Mollaret meningitis, increased levels of WBCs and PMNs in the CSF at the acute phase create speculative evidence that an epidermoid cyst may be the cause of this disorder.

The similarities in the CSF profile between epidermoid cyst–associated Mollaret meningitis and bacterial meningitis make the differential diagnosis difficult in the acute phase. In addition, it has been reported that intraspinal epidermoid cysts in children frequently cause recurrent episodes of bacterial meningitis associated with a dermal sinus tract. The sinus provides a pathway for bacterial entry into the subarachnoid space.1 However, no such sinus tract was found in our patient, and no bacterium was detected in the cyst liquid, blood, or CSF through culture and Gram-staining. Therefore, aseptic recurrent meningitis in our case is etiologically attributed to an intraspinal epidermoid cyst. This statement is further supported by the cessation of the illness after removal of the cyst from this patient.

The mechanisms by which an intraspinal epidermoid cyst forms or triggers meningitis are unclear. Intracranial
Epidermoid cysts are usually congenital and arise from ectodermal cell rests. It remains unclear whether intraspinal epidermoid cysts are also congenitally originated. In addition, it has been reported that lumbar puncture can result in the formation of spinal epidermoid cysts. However, this is not likely the case for our patient, because the cyst was situated ventral to the spinal cord, where it is unlikely to have been caused by puncture. In addition, the patient had no history of receiving a lumbar puncture before the symptoms. As for the pathogenesis of epidermoid cyst–associated meningitis, a proposed hypothesis suggests that the spontaneous rupture of the epidermoid cyst, on repeated occasions, may release contents into the subarachnoid space, causing relapsing chemical meningitis. The wall of the cyst might be closed and the cyst replenished after the rupture; we observed an intact wall around the cyst during surgery, when the patient was asymptomatic (Fig 3A).

Mollaret meningitis has been considered a “self-limited disease.” Unfortunately, the self-limited trend was not observed in the case reported here. On the contrary, the frequency of attacks increased over time. The symptom-free periods decreased from 2 to 6 months in the

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Source</th>
<th>Age</th>
<th>Gender</th>
<th>No. of Episodes</th>
<th>Histopathology</th>
<th>WBC, (\times 10^6/L)</th>
<th>Neutrophils, %</th>
<th>Protein, mg/dL</th>
<th>Glucose, mmol/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Current case</td>
<td>8 y</td>
<td>F</td>
<td>37</td>
<td>Epidermoid cyst</td>
<td>2520</td>
<td>99</td>
<td>103</td>
<td>2.0</td>
</tr>
<tr>
<td>2</td>
<td>Ref 11</td>
<td>33 y</td>
<td>F</td>
<td>NA</td>
<td>Epidermoid cyst</td>
<td>NA</td>
<td>75</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>3</td>
<td>Ref 11</td>
<td>34 y</td>
<td>M</td>
<td>NA</td>
<td>Epidermoid cyst</td>
<td>NA</td>
<td>75</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>4</td>
<td>Ref 12</td>
<td>36 y</td>
<td>F</td>
<td>NA</td>
<td>Epidermoid cyst</td>
<td>NA</td>
<td>60</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>5</td>
<td>Ref 10</td>
<td>7 y</td>
<td>M</td>
<td>5</td>
<td>Epidermoid cyst</td>
<td>4000</td>
<td>95</td>
<td>67</td>
<td>1.83</td>
</tr>
<tr>
<td>6</td>
<td>Ref 13</td>
<td>5 y</td>
<td>M</td>
<td>2</td>
<td>Epidermoid cyst</td>
<td>700</td>
<td>53</td>
<td>426</td>
<td>4.83</td>
</tr>
<tr>
<td>7</td>
<td>Ref 14</td>
<td>32 y</td>
<td>M</td>
<td>30</td>
<td>Epidermoid cyst</td>
<td>4300</td>
<td>89</td>
<td>139</td>
<td>2.50</td>
</tr>
<tr>
<td>8</td>
<td>Ref 1</td>
<td>44 y</td>
<td>M</td>
<td>2</td>
<td>Epidermoid cyst</td>
<td>4200</td>
<td>97</td>
<td>123</td>
<td>2.3</td>
</tr>
<tr>
<td>9</td>
<td>Ref 15</td>
<td>25 mo</td>
<td>F</td>
<td>3</td>
<td>Epidermoid cyst</td>
<td>5000</td>
<td>100</td>
<td>160</td>
<td>3.28</td>
</tr>
<tr>
<td>10</td>
<td>Ref 16</td>
<td>34 y</td>
<td>M</td>
<td>30</td>
<td>Epidermoid cyst</td>
<td>4625</td>
<td>89</td>
<td>139</td>
<td>2.55</td>
</tr>
<tr>
<td>11</td>
<td>Ref 17</td>
<td>21 y</td>
<td>F</td>
<td>2</td>
<td>Epidermoid cyst</td>
<td>2810</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>12</td>
<td>Ref 17</td>
<td>16 y</td>
<td>F</td>
<td>NA</td>
<td>Epidermoid cyst</td>
<td>NA</td>
<td>75</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>13</td>
<td>Ref 17</td>
<td>14 mo</td>
<td>F</td>
<td>4</td>
<td>Epidermoid cyst</td>
<td>1120</td>
<td>89</td>
<td>40</td>
<td>3.45</td>
</tr>
</tbody>
</table>

NA indicates not available; F, female; M, male.

<table>
<thead>
<tr>
<th>Case No.</th>
<th>Reference No.</th>
<th>Age, y</th>
<th>Gender</th>
<th>No. of Episodes</th>
<th>HSV PCR</th>
<th>WBC, (\times 10^6/L)</th>
<th>Neutrophils, %</th>
<th>Protein, mg/ dl</th>
<th>Glucose, mmol/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4 39 F 6 P</td>
<td>20</td>
<td>0</td>
<td>70</td>
<td>2.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>18 31 F 2 P</td>
<td>267</td>
<td>9</td>
<td>33.3</td>
<td>4.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>18 55 M 2 P</td>
<td>538</td>
<td>5</td>
<td>1.11</td>
<td>4.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>19 33 M 6 P</td>
<td>457</td>
<td>NA</td>
<td>118</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>20 35 F 3 P</td>
<td>360</td>
<td>15</td>
<td>49</td>
<td>4.67</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>21 32 F 8 P</td>
<td>470</td>
<td>&lt;10</td>
<td>250</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>21 33 F 2 P</td>
<td>70</td>
<td>&lt;10</td>
<td>110</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>21 56 F 3 P</td>
<td>700</td>
<td>&lt;10</td>
<td>170</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>21 40 F 3 P</td>
<td>310</td>
<td>&lt;10</td>
<td>250</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>21 33 F 2 P</td>
<td>200</td>
<td>&lt;10</td>
<td>120</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>21 52 F 2 P</td>
<td>1380</td>
<td>&lt;10</td>
<td>360</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>21 44 M 2 P</td>
<td>75</td>
<td>&lt;10</td>
<td>183</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>21 38 F 3 P</td>
<td>360</td>
<td>&lt;10</td>
<td>151</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>21 35 M 2 P</td>
<td>185</td>
<td>18</td>
<td>620</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>22 41 F NA P</td>
<td>144</td>
<td>0</td>
<td>750</td>
<td>2.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>5 23 F 3 P</td>
<td>233</td>
<td>NA</td>
<td>89.7</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>5 36 F 3 P</td>
<td>735</td>
<td>NA</td>
<td>155</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>5 37 F 3 P</td>
<td>138</td>
<td>NA</td>
<td>63.7</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>5 32 F 2 P</td>
<td>242</td>
<td>NA</td>
<td>5544</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>24 27 M 3 P</td>
<td>868</td>
<td>6</td>
<td>81</td>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>25 44 F 3 P</td>
<td>433</td>
<td>NA</td>
<td>144</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>25 40 M 7 P</td>
<td>736</td>
<td>13</td>
<td>184</td>
<td>3.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NA indicates not available; F, female; M, male; P, positive; HSV, herpes simplex virus; PCR, polymerase chain reaction.
early period to 3 to 7 days in the last 3 months of her illness. One possible explanation for this increasing frequency is that repeated cyst rupture causes an incomplete repair of the cyst wall, which becomes thinner over time and thereby ruptures more easily. Thus, it seems that so called self-limited Mollaret meningitis may not be a clinical feature of the disease as caused by the epidermoid cyst.

CONCLUSIONS

Our clinical data, laboratory results, and imaging analysis support intraspinal epidermoid cyst as one of the etiologies of Mollaret meningitis. Intraspinal and intracranial epidermoid cyst–caused Mollaret meningitis have similar clinical manifestations and CSF features during an attack. Because of the importance of surgical treatment for this disease, spinal neuroimaging studies (computed tomography or MRI) should be performed for patients suspected of having epidermoid cyst–caused Mollaret meningitis when no lesion has been observed in the intracranial contents.

REFERENCES

Human Milk Intake and Retinopathy of Prematurity in Extremely Low Birth Weight Infants

Cherrie D. Heller, MD, MPH*, Michael O’Shea, MD, MPH*, Qing Yao, PhD*, John Langer, MSc*, Richard A. Ehrenkranz, MD*, Dale L. Phelps, MD*, W. Kenneth Poole, PhD*, Barbara Stoll, MD*, Shahnaz Duara, MD*, William Oh, MD*, James Lemons, MD*, Brenda Poindexter, MD, MSc†, for the NICHD Neonatal Research Network

*Department of Pediatrics, Wake Forest University School of Medicine, Winston-Salem, North Carolina; †RTI International, Research Triangle Park, North Carolina; Department of Pediatrics, Yale University School of Medicine, New Haven, Connecticut; Department of Pediatrics, University of Rochester School of Medicine and Dentistry, Rochester, New York; Department of Pediatrics, Emory School of Medicine, Atlanta, Georgia; Department of Pediatrics, University of Miami School of Medicine, Miami, Florida; Department of Pediatrics, Women and Infants Hospital, Providence, Rhode Island; Department of Pediatrics, Indiana University School of Medicine, Indianapolis, Indiana; National Institute of Child Health and Human Development, Rockville, Maryland.

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

ABSTRACT

OBJECTIVES. Our goal was to analyze the association between human milk intake and severe retinopathy of prematurity in extremely low birth weight infants.

PATIENTS AND METHODS. This study is a secondary analysis of data collected for a trial of glutamine supplementation in extremely low birth weight infants (birth weight <1000 g). Among the 1433 participants in that trial, data are available regarding human milk intake and the occurrence of severe retinopathy of prematurity (defined in this study as retinopathy of prematurity treated surgically) for 1057 infants. The volume of human milk intake was expressed as the mean volume (milliliters per kilogram per day) and the mean proportional volume (proportion of total nutritional intake) from birth to discharge or transfer. Using logistic regression, we estimated odds ratios and 95% confidence intervals for any human milk intake and, among infants who received human milk, for each 10 mL/kg per day and each 10% increase in volume.

RESULTS. Of the 1057 infants included in this cohort, 788 infants (75%) received at least some human milk. Among these milk-fed infants, the median volume of human milk intake was 30 mL/kg per day (interquartile range: 6–83 mL/kg per day), and the median proportional volume of human milk intake was 0.18 (interquartile range: 0.03–0.66). One hundred sixty-three infants (15%) developed severe retinopathy of prematurity.

CONCLUSIONS. Human milk intake was not associated with a decreased risk of severe retinopathy of prematurity.
Retinopathy of Prematurity (ROP) is a vascular disorder of the retina affecting infants born prematurely. The reported incidence of severe ROP (stage 3) in infants born <1000 g varies from 14% to 40%,\(^1\) representing 96% of all cases of stage 3 ROP. After cryotherapy\(^4\) or laser surgery,\(^5\) vision is severely impaired in 44% of children and 18% of eyes, respectively. In addition, myopia, strabismus, and amblyopia occur frequently.\(^6\)\(^,\)\(^7\) The pathogenesis of ROP is incompletely understood, but factors that have been implicated include exposure of the developing retina to abnormal oxygen levels\(^8\) and cytotoxic reactive oxygen species,\(^9\) the premature infants’ reduced antioxidant defenses,\(^10\)\(^,\)\(^11\) and decreased ability to synthesize long chain polyunsaturated fatty acids (LCPUFAs).\(^12\)\(^-\)\(^14\)

Because it contains LCPUFAs\(^15\)\(^,\)\(^16\) and antioxidant enzymes, human milk (HM)\(^17\)\(^-\)\(^19\) might influence the development of ROP. The objective of this study was to analyze the association between HM intake and the development of severe ROP, defined in this study as ROP that was treated surgically. We hypothesized that infants fed HM would be less likely to develop severe ROP and that higher levels of HM intake would be associated with a lower incidence of severe ROP.

**PATIENTS AND METHODS**

**Sample**

The cohort for this observational study was derived from the 1433 infants enrolled in a multicenter, randomized clinical trial of parenteral glutamine supplementation, conducted by the National Institute of Child Health and Human Development (NICHD) Neonatal Research Network (NRN) from October 1999 to September 2001.\(^20\) The trial included infants born at 14 study centers with birth weights between 401 and 1000 g. Excluded from the trial were infants with major congenital anomalies, congenital nonbacterial infection, or hypoxia with bradycardia for >2 hours or blood pH <6.80, and those for whom full medical support was not provided. For our study, we also excluded 104 infants who were never enterally fed (Table 1).

**Data Sources**

All data, except those about HM intake, were obtained from the NRN Generic Database or the NRN 18–22 Month Follow up Study database, for which data are collected from medical charts by research nurses using prespecified definitions as described in previous publications.\(^21\)\(^,\)\(^22\) Data about HM intake were collected as part of the aforementioned clinical trial of parenteral glutamine.\(^20\) During that trial, the type and volume of nutritional intake was recorded for all subjects from the day of birth until the earliest of the following events: discharge, transfer, death, or the 120th day of life. Data were recorded daily until the infant reached full enteral feedings. Once an infant had been fully enterally fed for 72 hours, data collection was reduced to 3 days per week: Monday, Wednesday, and Friday (MWF). If enteral feedings were discontinued or parenteral supplementation was reintroduced, daily data collection was

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Infants With Complete Data (N = 1057)</th>
<th>Excluded Infants (N = 104)*</th>
<th>Infants With Missing Data for the Outcome Died (N = 126)</th>
<th>No Exam or Follow-up Data (N = 146)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prenatal care</td>
<td>93</td>
<td>91</td>
<td>94</td>
<td>87</td>
</tr>
<tr>
<td>Antenatal steroids</td>
<td>81</td>
<td>61</td>
<td>78</td>
<td>77</td>
</tr>
<tr>
<td>Maternal hypertension</td>
<td>27</td>
<td>18</td>
<td>22</td>
<td>19</td>
</tr>
<tr>
<td>Vaginal delivery</td>
<td>38</td>
<td>49</td>
<td>52</td>
<td>45</td>
</tr>
<tr>
<td>Duration of rupture of membranes, h</td>
<td>41 (136)</td>
<td>15 (49)</td>
<td>24 (59)</td>
<td>52 (151)</td>
</tr>
<tr>
<td>Male gender</td>
<td>45</td>
<td>57</td>
<td>53</td>
<td>55</td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Black</td>
<td>45</td>
<td>50</td>
<td>46</td>
</tr>
<tr>
<td>White</td>
<td>39</td>
<td>28</td>
<td>33</td>
<td>37</td>
</tr>
<tr>
<td>Other</td>
<td>16</td>
<td>22</td>
<td>21</td>
<td>13</td>
</tr>
<tr>
<td>Gestational age, wk</td>
<td>26.2 (2)</td>
<td>24.6 (2)</td>
<td>25 (1.7)</td>
<td>26.3 (2.2)</td>
</tr>
<tr>
<td>Outborn</td>
<td>6</td>
<td>14</td>
<td>10</td>
<td>7</td>
</tr>
<tr>
<td>S-min Apgar score of &lt;6</td>
<td>20</td>
<td>42</td>
<td>33</td>
<td>19</td>
</tr>
<tr>
<td>Multiple birth</td>
<td>21</td>
<td>27</td>
<td>27</td>
<td>19</td>
</tr>
<tr>
<td>Birth weight, g</td>
<td>785 (131)</td>
<td>660 (135)</td>
<td>693 (135)</td>
<td>795 (148)</td>
</tr>
<tr>
<td>FGR-10*</td>
<td>1.24 (0.3)</td>
<td>1.26 (0.3)</td>
<td>1.25 (0.2)</td>
<td>1.23 (0.3)</td>
</tr>
<tr>
<td>Respiratory distress syndrome</td>
<td>97</td>
<td>100</td>
<td>97</td>
<td>96</td>
</tr>
<tr>
<td>Pneumothorax</td>
<td>5</td>
<td>26</td>
<td>23</td>
<td>12</td>
</tr>
<tr>
<td>Day of first feeding</td>
<td>7.6 (6.5)</td>
<td>NA</td>
<td>9.2 (8.1)</td>
<td>8.1 (7.3)</td>
</tr>
<tr>
<td>HM (yes)</td>
<td>75</td>
<td>0</td>
<td>76</td>
<td>63</td>
</tr>
</tbody>
</table>

Data are column percentages of infants with the attribute or the mean (SD) for infants with the attribute. NA indicates not applicable.

* Excluded infants were those who were never enterally fed.

* FGR-10 indicates birth weight/10th percentile of birth weight for gestational age according to gender (using growth curves described by Alexander et al\(^23\)\).
Definitions

**HM Intake**

HM intake was defined in 3 ways. In analyses involving the entire cohort, HM intake was defined as a dichotomous variable (at least some HM or no HM feedings). In analyses involving only those infants who received HM, the magnitude of the intake was expressed in 2 ways: (1) the mean volume of HM intake, expressed as milliliters per kilogram per day, from birth to discharge or transfer, and as (2) the proportional volume of all recorded nutritional fluid intake (parenteral nutrition and enteral nutrition, but not other glucose-containing solutions) that consisted of HM, from birth through discharge or transfer. At the time this study was conducted, preterm formulas did not contain supplemental omega-3 LCPUFAs, and none of the infants received donor milk.

Because this study was an observational study using previously collected data, no efforts were made to influence the HM intake or any other facet of nutritional management of the infants in this study.

**ROP Surgery**

Surgical therapy for ROP was used as a surrogate for severe or threshold ROP because the data set available to us did not contain information about the number of “clock hours” of involvement, which is needed to classify ROP as “threshold,” “prethreshold,” or “less than prethreshold” ROP. The data set included information about whether surgery was performed for ROP for all infants through discharge or transfer (whichever occurred first), and whether ROP surgery occurred beyond this time, for those infants who returned for a follow-up clinic visit at 18 and/or 30 months’ adjusted age. Infants were classified as to whether they had developed ROP that was treated surgically. However, infants who did not undergo ROP surgery before discharge or transfer and did not return for follow-up at either 18 or 30 months’ adjusted age were classified as “lost to follow-up,” meaning that data about the outcome of interest (ROP treated surgically) were missing for these infants. In addition, data were not included in the analysis for infants who died before reaching an age at which they could be classified by ROP outcome.

Although this study did not determine the criteria for performing ROP surgery, these criteria are generally well accepted and based on those defined in the CRYO-ROP study.

**Potential Confounders**

In the NRN Generic database, prenatal care is defined as ≥1 prenatal care visits, and antenatal steroids as maternal treatment with corticosteroids before delivery. Mothers are classified as having hypertension if ≥1 of the following is listed in the mother’s chart: diagnosis of hypertension, systolic blood pressure >140 mm Hg, or a diastolic blood pressure >90 mm Hg. Maternal education was categorized as less than high school, high school, or greater than high school. For this study, early onset sepsis (suspected or confirmed) is defined as treatment with antibiotics for ≥5 days, beginning in the first 72 hours after birth, regardless of blood culture results. An infant was classified as having respiratory distress syndrome if he/she had both “clinical features of respiratory distress” and an “abnormal chest radiograph within 24 hours of birth.” For the current study, the fetal growth ratio-10 (FGR) was used as an indicator of fetal growth. This was determined by dividing the birth weight of the infant by the 10th percentile of birth weight for the appropriate gestational age and gender from growth charts developed by Alexander et al. The FGR is a more informative indicator of fetal growth than dichotomous measures, such as birth weight less than the 10th percentile. Although other investigators based the FGR on the gestational age–specific median birth weight, the reference values provided by Alexander et al do not include gestational age– and gender-specific medians but do contain these values for the 10th percentiles of birth weight. Therefore, we standardized the birth weights in this study to the 10th percentiles of these curves.

**Statistical Analysis**

Our analytical approach was to test whether any HM intake was associated with severe ROP and then to look at whether the level of HM intake was associated with severe ROP among those subjects who received HM by using 2 continuous measures of HM intake. Unadjusted
odds ratios (ORs) and 95% confidence intervals (CIs) were estimated for the association of HM intake (any versus none, mean volume and proportional volume) and severe ROP.

**Selection of Potential Confounding Variables**
Factors that were considered as potential confounding variables met the following criteria: (1) previous studies suggested that the factor was associated with either HM or ROP, (2) the factor was ascertained before the feedings were initiated, so it could not be involved in a pathway linking HM intake and ROP, and (3) data were available for the factor in the NRN Generic database. For factors that met these criteria, we analyzed associations between the factor and severe ROP and between the factor and HM. If the P values for both of these associations were ≤0.2, the factor was regarded as a potentially confounding variable and was included in the multivariate analyses, as described below.

**Specification of Multivariate Regression Models**
The potential confounding factors, identified in the manner described above, were classified as either prenatal or postnatal factors, according to the time of their occurrence (factors listed in Table 2). In the first step of multivariate analysis, HM and the prenatal factors that were identified as potential confounders were entered into a logistic model as independent variables with ROP (no ROP/ROP not requiring surgery versus ROP treated with surgery) as the outcome variable. The prenatal variables that were found to be independently associated with ROP outcome, at a significance level of *P* < .1, were retained in the model. In the second step, the postnatal variables that were identified as potential confounders were entered into the model. Then, those postnatal variables that were not found to be independently associated with ROP outcome at a significance level of *P* < .1 were eliminated from the model. In the final step of multivariate analysis, ORs and 95% CIs were estimated and adjusted for the retained potential confounding factors. For the analyses in which HM intake was expressed as the mean volume or the proportional volume, ORs were reported for each 10 mL/kg per day and each 10% increase in intake, respectively. The generalized estimating equation method was used to account for the inclusion of infants born of multiple gestations (twins and triplets).

**Analysis of Possible Effect Modifiers**
To explore whether certain factors influence the strength of association between HM and the severity of ROP (ie, effect modification) we estimated ORs for strata, which were defined in terms of 4 factors that previous research suggests might be effect modifiers:

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Associations Between Perinatal Factors HM and Associations Between Perinatal Factors and ROP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
<td>No HM (N = 353)</td>
</tr>
<tr>
<td>Prenatal care</td>
<td>83</td>
</tr>
<tr>
<td>Antenatal steroids</td>
<td>72</td>
</tr>
<tr>
<td>Labor</td>
<td>68</td>
</tr>
<tr>
<td>Maternal hypertension</td>
<td>25</td>
</tr>
<tr>
<td>Vaginal delivery</td>
<td>36</td>
</tr>
<tr>
<td>Duration of rupture of membranes, h</td>
<td>28 (76)</td>
</tr>
<tr>
<td>Maternal education</td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>34</td>
</tr>
<tr>
<td>High school</td>
<td>42</td>
</tr>
<tr>
<td>More than high school</td>
<td>23</td>
</tr>
<tr>
<td>Male gender</td>
<td>48</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>64</td>
</tr>
<tr>
<td>White</td>
<td>27</td>
</tr>
<tr>
<td>Other</td>
<td>9</td>
</tr>
<tr>
<td>Gestational age, wk</td>
<td>26.2 (2)</td>
</tr>
<tr>
<td>Outborn</td>
<td>7</td>
</tr>
<tr>
<td>5-min Apgar score of &lt;6</td>
<td>23</td>
</tr>
<tr>
<td>Multiple birth</td>
<td>20</td>
</tr>
<tr>
<td>Birth weight, g</td>
<td>783 (140)</td>
</tr>
<tr>
<td>FGR-10</td>
<td>1.2 (0.2)</td>
</tr>
<tr>
<td>Respiratory distress syndrome</td>
<td>97</td>
</tr>
<tr>
<td>Pneumothorax</td>
<td>6</td>
</tr>
<tr>
<td>Treatment for early onset sepsisb</td>
<td>54</td>
</tr>
<tr>
<td>Day of first feeding</td>
<td>7.7 (6.7)</td>
</tr>
</tbody>
</table>

Data are column percentages (ie, percentages of infants with and without HM intake, or with and without ROP treated surgically who had the listed attribute) or the mean (SD) for infant with the attribute. *P* values are for a *χ²* test of the association between the attribute and either HM intake or severe ROP treated surgically.

*FGR-10 indicates birth weight/10th percentile of birth weight for gestational age according to gender (using growth curves by Alexander et al)*.

* Suspected or confirmed: defined as treatment with antibiotics for at least 5 days, beginning in the first 72 hours after birth.
maternal educational level (a surrogate measure of socioeconomic status), ethnicity, FGR-10 (as a measure of the adequacy of fetal growth), and postmenstrual age at first feeding (as an indicator of the timing of initial exposure to HM).

**Power Calculation**

Given the estimated SEs of the ORs from this study for the HM (any or none) and HM (mL/kg per day), the available sample sizes were sufficient to detect a reduction in the OR to \( \leq 0.52 \) for HM (any or none) and to \( \leq 0.92 \) for HM (mL/kg per day) with a power of 80% and a 2-tailed significance level of .05. The OR for HM (mL/kg per day) is the reduction in the OR per 10 mL/kg per day.

**RESULTS**

**Sample**

Of 1433 infants enrolled in the NRN trial of glutamine supplementation, 1329 infants were enterally fed. Of these, 272 had missing data for the outcome (126 died, 5 did not have an eye examination, and 141 did not come for follow-up at 18 or 30 months’ adjusted age). Thus, there were 1057 infants with data for both the exposure and outcome of interest (Table 1). The mean (SD) day of life and mean (SD) postmenstrual age at the time of discharge or transfer were 87.9 (25) days and 38.7 (2.9) weeks, respectively. Eighty-three (7%) infants were transferred to another hospital before discharge home at a mean (SD) day of life and mean (SD) postmenstrual age at transfer of 55.5 (26.3) days and 34.5 (3.1) weeks, respectively.

**Description of ROP and HM Feeding**

Of 1057 infants for whom complete data were available, 163 (15%) developed severe ROP. Seven hundred eighty-eight infants (75%) received HM. Among these infants, the median (interquartile range) of the volume of HM intake was 30 (6–83) mL/kg per day and the median (interquartile range) for the proportion of total nutritional fluids (enteral nutrition plus parenteral nutrition) consisting of HM was 0.18 (0.03–0.66).

The unadjusted OR (95% CI) for the association between any HM intake and severe ROP was 1.41 (0.94–2.13; \( P = .1 \)). Among infants who received any HM, increasing volume of intake (mL/kg per day) was associated with a decreased unadjusted odds of developing severe ROP (OR [95% CI] per 10 mL/kg per day increase in intake: \( 0.95 \) [0.91–1.00]; \( P = .05 \)). The unadjusted OR (95% CI) for developing severe ROP with increasing proportional volume of HM intake per 10% increase was \( 0.96 \) (0.91–1.02; \( P = .2 \)).

Irrespective of whether HM intake was expressed as a dichotomy (any HM intake versus none) or as a continuous variable (mean volume of intake or mean proportional volume), the following factors were identified as potential confounders: antenatal steroids, maternal education, ethnicity, mode of delivery, gestational age, treatment for proven or suspected early onset sepsis, and pneumothorax (Table 2). In addition, the following factors were identified as potential confounders when HM intake was expressed as a continuous variable: occurrence of labor, maternal hypertension, outborn status, low 5-minute Apgar score, and multiple birth. Lastly, infant gender and receipt of surfactant were identified as potential confounders only when HM intake was express as mean volume. The group to which infants were randomized in the glutamine supplementation trial (from which the nutritional data for this study were obtained) did not influence either HM intake or ROP outcome. The diagnosis of patent ductus arteriosus or ultrasonographic evidence of severe intraventricular hemorrhage or white matter injury were not associated with HM intake and, therefore, were not included as confounders.

Although variables associated with chronic lung disease (eg, duration of mechanical ventilation and supplemental oxygen, and receipt of postnatal steroids) were associated with ROP and HM intake in this study, they were not included as confounders in the analysis because they are likely intermediates in the presumed pathway linking HM and the development of ROP.

ORs for HM, and for the variables which, in multivariate models, had a statistically significant association with severe ROP, are depicted in Figs 1 and 2. The adjusted ORs (95% CI) for developing severe ROP for each of the models containing the 3 HM exposure vari-
The strength of the association between HM and ROP was not influenced by maternal education, ethnicity, FGR-10, or postmenstrual age at first enteral feeding. In addition, analyses in which "study center" was included as a potential confounder did not produce significantly different results from those in which this variable was not included; therefore, this variable was not retained in the final regression models.

DISCUSSION

Our study was undertaken to investigate whether 1 or more constituents of HM are protective against severe ROP. Our hypothesis was based on the fact that HM contains immunomodulatory substances, such as secretory immunoglobulin A, lactoferrin, lysozyme, cytokines, oligosaccharides, antioxidant enzymes and cellular components. These factors are thought to influence immune defenses of the infant, which may explain the lower risk of necrotizing enterocolitis and sepsis among HM-fed infants. HM also contains docosahexaenoic acid (DHA), which has an important role in the developing brain and retina. Contrary to our hypotheses, in the sample of extremely low birth weight infants whom we studied, neither receipt of HM nor increasing intake of HM were associated with a decreased risk of developing severe ROP.

The point estimate for the OR for any HM (versus none) was 1.47 (0.94–2.32), P = .09 (Fig 1); proportional volume of HM intake: 0.98 (0.91–1.05), P = .52 (Fig 2); and mean volume of HM intake (mL/kg per day) was 0.98 (0.92–1.04), P = .43 (not shown). The variables included in the model containing HM intake by mean volume were pneumothorax, ethnicity, maternal hypertension, day to first enteral feeding, antenatal steroids, gestational age at birth, and maternal education.

There was no evidence of a statistically significant dose response relationship between increasing HM intake and development surgical ROP or a threshold of HM intake above which the probability of developing severe ROP declined (Fig 3).
ined previously, with conflicting results. Hylander et al\(^1\) found that HM (any versus none) was associated with a decreased risk of ROP, even after controlling for gestational age, duration of supplemental oxygen therapy, 5-minute Apgar score, and ethnicity (OR: 0.46; 95% CI: 0.19–0.93), but no dose response was detected with increasing levels of HM intake; no association between HM and ROP was found in a study by Furman et al.\(^2\).

Schanler et al\(^3\) reported that infants who received only HM, compared with those who were fed HM in addition to either donor HM or premature formula, were at lower risk for ROP, suggesting that higher intake of HM is associated with a lower risk of ROP.

It would be difficult to conduct a randomized, controlled trial of HM using infants’ own mother’s milk. Observational studies such as this one may be confounded by factors that influence the mother’s decision to provide HM and the length of time mothers choose to provide HM, such as socioeconomic status, ethnicity, maternal health status, and infant health status. If these same factors also influence the risk of ROP, inconsistency in the results of studies of HM and ROP risk could be attributable, at least in part, to variation in the extent to which confounding has been controlled.

Other possible sources of variation across studies of HM and ROP include the amount of HM fed to infants, the age at which enteral feedings are initiated, the rapidity with which feedings are advanced, and the composition (eg, DHA content) of the HM. The DHA content of HM from women in the United States is very low, presumably because of low fish intake.\(^4\) In addition, in the present study, intake of HM-fed infants was relatively low, comprising only ~15% of their total nutrition throughout their hospitalization. Although this is a very low level of intake, it is likely representative of the current intake of HM in extremely low birth weight infants in the United States. Also in this study, feedings were initiated after the first week of life in a majority of study infants. If HM does lower the risk of severe ROP, it is more likely to be effective when the antiinflammatory and antioxidant components of HM are provided earlier in life during a time when infants are exposed to high levels of oxidative stress and inflammatory cytokines. Averaging HM feedings over the entire hospitalization, as was done in this study, may not be sensitive enough to detect a potential effect of HM on ROP that operates over a specific limited time span.

Three limitations of our study that could have masked an association of HM and ROP should be noted. First, because the majority of infants were discharged before 42 weeks’ postmenstrual age and no data were collected about the results of eye examinations performed after discharge, we were not able to fully classify infants as to the development of ROP or its degree of severity, unless they developed disease severe enough to be treated surgically. Combining infants with no ROP and infants with ROP not treated surgically would be expected to attenuate an association between the level of HM intake and ROP, if such an association exists. It may be that HM intake is protective against the development of less severe forms of ROP. Second, data were missing for slightly >10% of the sample who did not require surgery for ROP before discharge and did not return for follow-up visits at 18 or 30 months’ corrected age. Infants with missing data for ROP outcome were less likely to have received HM and more likely to have had a pneumothorax. The latter was associated with a higher risk of severe ROP, suggesting that the bias resulting from losing infants to follow-up after discharge would have attenuated a potential “protective” effect of HM. A third limitation of this study is the absence of defined criteria for performing surgery for ROP across the study centers. However, we feel this is largely a theoretical concern as criteria for performing peripheral retinal ablation surgery are generally well accepted and based on those defined in the CRYO-ROP study.\(^1\)

CONCLUSIONS

Despite our finding of no association between HM and severe ROP, HM has been found to have several other associated benefits, including a reduced risk of late onset infection and necrotizing enterocolitis,\(^3\) improved tolerance of enteral feedings,\(^2\) and possible beneficial effect on neurodevelopmental outcomes.\(^3\) Future research should be directed toward investigating the association between HM and ROP in an exclusively HM-fed group with complete follow-up of infants after discharge to at least 42 weeks’ postmenstrual age. Lactation support and initiation of enteral feedings as early as possible could be used to maximize HM intake in HM-fed infants.

ACKNOWLEDGMENTS

We thank the children and their parents who participated in the randomized trial of glutamine supplementation and made this study possible.

Members of the NICHD Neonatal Research Network Trial (and NICHD grants) include University of Cincinnati: Alan Jobe, MD, PhD (chairman and principal investigator); University of California at San Diego (U10 HD40461): Neil N. Finer, MD (principal investigator), Yvonne Vaucher, MD (follow-up principal investigator), Chris Henderson, and Martha Fuller; Case Western Reserve University (U10 HD21364): Avroy A. Fanaroff, MB, BCh (principal investigator), Dee Wilson, MD (follow-up principal investigator), Nancy Newman, RN, and Bonnie Siner, RN; University of Cincinnati (U10 HD27853, M01 RR 08084): Vivek Narendran, MD, MRCP, Jean Steichen, MD (follow-up principal investigator), Marcia Mersmann, RN, and Teresa Gratton, RN; Emory University (U10 HD27851): Barbara J. Stoll, MD (principal investigator), Ira Adams-Chapman, MD (follow-up principal investigator), and Ellen Hale, RN; Indi-
Ana University (U10 HD27856, M01 RR 00750): James A. Lemons, MD (principal investigator), Anna Dusick, MD (follow-up principal investigator), Lucy Miller, RN, and Leslie Richard, RN; University of Miami (U10 HD21397): Charles R. Bauer, MD (principal investigator), Shahnaz Duara, MD, and Ruth Everett, RN; National Institute of Child Health and Human Development: Linda L. Wright, MD, Rosemary Higgins, MD, and James Hanson, MD; University of New Mexico (U10 HD27881, M01 RR00977): Lu-Ann Papile, MD (principal investigator), and Conra Backstrom, RN; Research Triangle Institute (U01 HD36790): W. Kenneth Poole, PhD, Abhik Das, PhD, Betty Hastings, and Carolyn Petrie, MS; Stanford University (U10 HD27880, M01 RR 00070): David K. Stevenson, MD (principal investigator), Susan Hintz, MD (follow-up principal investigator), and Bethany Ball, BS; University of Tennessee at Memphis (U10 HD21415): Sheldon B. Korones, MD (principal investigator), Henrietta Bada, MD, and Tina Hudson, RN; University of Texas Health Science Center at Houston (U10 HD21373): Jon E. Tyson, MD, MPH (principal investigator), and Georgia McDavis, RN; University of Texas Southwestern Medical Center (U10 HD40689): Abbot R. Laptook, MD (principal investigator), Roy Heyne, MD, Susie Madison, RN, and Janet Morgan, RN; Wayne State University (U10 HD21385): Seetha Shankaran, MD (principal investigator), Yvette Johnson, MD, Geraldine Muran, RN, and Debbie Kennedy, RN; Women and Infants Hospital (U10 HD27904): William Oh, MD (principal investigator), Betty Vohr, MD, Angelita Hensman, RN, and Lucy Noel, RN; Yale University (U10 HD27871, M01 RR 06022): Richard A. Ehrenkranz, MD (principal investigator and follow-up investigator), Patricia Gettner, RN, and Elaine Romano, RN; and University of Alabama at Birmingham (U10 HD34216): Waldemar A. Carlo, MD (principal investigator), Myriam Peralta, MD, Monica V. Collins, RN, and Vivien Phillips, RN.

REFERENCES

FETAL PULSE OXIMETRY AND CESAREAN DELIVERY

"Background: Knowledge of fetal oxygen saturation, as an adjunct to electronic fetal monitoring, may be associated with a significant change in the rate of cesarean deliveries or the infant’s condition at birth.  
Methods: We randomly assigned 5341 multiparous women who were at term and in early labor to either ‘open’ or ‘masked’ fetal pulse oximetry. In the open group, fetal oxygen saturation values were displayed to the clinician. In the masked group, the fetal oxygen sensor was inserted, and the values were recorded by computer, but the data were hidden. Labor complicated by a non-reassuring fetal heart rate before randomization was documented for subsequent analysis, . . .  
Conclusions: Knowledge of the fetal oxygen saturation is not associated with a reduction in the rate of cesarean delivery or with improvement in the condition of the newborn. (ClinicalTrials.gov number, NCT00098709.)"  
Noted by JFL, MD
Malpractice Claims Involving Pediatricians: Epidemiology and Etiology

Aaron E. Carroll, MD, MS¹, Jennifer L. Buddenbaum, MHA²

¹Children’s Health Services Research, Indiana University School of Medicine, Indianapolis, Indiana; ²Regenstrief Institute for Health Care, Indianapolis, Indiana

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

ABSTRACT

OBJECTIVE. Our goals were to examine malpractice claims data that are specific to the specialty of pediatrics and to provide a better understanding of the effect that malpractice has on this specialty.

METHODS. The Physician Insurers Association of America is a trade association of medical malpractice insurance companies. The data contained in its data-sharing project represent ~25% of the medical malpractice claims in the United States at a given time. Although this database is not universally comprehensive, it does contain information not available in the National Practitioner Data Bank, such as information on claims that are not ultimately paid and specialty of the defendant. We asked the Physician Insurers Association of America to perform a query of its data-sharing project database to find malpractice claims reported between January 1, 1985, and December 31, 2005, in which the defendant’s medical specialty was coded as pediatrics. Comparison data were collected for 27 other specialties recorded in the database.

RESULTS. During a 20-year period (1985–2005), there were 214,226 closed claims reported to the Physician Insurers Association of America data-sharing project. Pediatricians account for 2.97% of these claims, making it 10th among the 28 specialties in terms of the number of closed claims. Pediatrics ranks 16th in terms of indemnity payment rate (28.13%), with dentistry ranked highest at 43.35%, followed by obstetrics and gynecology at 35.50%. Indemnity payment refers to settlements or awards made directly to plaintiffs as a result of claim-resolution process. Data are presented on changes over time, claim-adjudication status, expenses on claims, the causes of claims, and injuries sustained.

CONCLUSIONS. Malpractice is a serious issue. Some will read the results of this analysis and draw comfort; others will view the same data with alarm and surprise. Regardless of how one interprets these findings, they are important in truly informing the debate with generalizable facts.
The medical malpractice liability system in the United States is meant to serve 2 purposes: (1) to provide compensation to individuals who suffer negligent medical injuries, and (2) to deter professional negligence. Many feel that our current medical liability system is broken, pointing to issues such as rising medical malpractice premiums, untimely and inequitable compensation for truly injured patients, and the questionable impact that litigation has on health care quality, cost, and access to care. This has led to the emergence of medical malpractice as a hotly debated issue on the national health policy stage.

The American Medical Association (AMA) has deemed that America is in the midst of a medical liability crisis and is aggressively pursuing medical liability reform. In support of this position, they cite the rapid rise in jury awards that has occurred between 1997 and 2003 (increasing from a median award of $157,000 to $300,000) and the similar escalations in settlement amounts (increasing from a mean of $212,861 to $322,544). According to the General Accounting Office, losses on medical malpractice claims, which make up the largest part of insurers’ costs, seem to be the primary driver of malpractice insurance rate increases over the long run. These losses included both payments to plaintiffs to resolve claims and the costs associated with defending claims. The AMA believes that rising premiums are resulting in a patient access crisis as physicians are forced to limit services, retire early, or move to states where liability premiums are stable.

Data have demonstrated that certain specialties have been impacted more than others. For example, malpractice premiums for all physicians nationwide increased an average of 15% between 2000 and 2002, whereas specialties such as obstetrics and gynecology (22%), general surgery (33%), and internal medicine (33%) have seen higher increases. It is important that each specialty understand the etiology, epidemiology, and pathology of malpractice claims because it can provide insight into steps that can be taken to decrease the risk of such claims. Although many specialties have begun to explore their own data, to date only limited data have been presented regarding malpractice claims within the specialty of pediatrics. The majority of pediatric data published have related to subspecialties, such as pediatric emergency medicine, radiology, or neonatology.

The American Academy of Pediatrics’ Division of Health Policy has monitored pediatricians’ experiences with medical liability since 1987 through its Periodic Survey of Fellows. This survey is a self-administered questionnaire mailed to a random sample of 1600 active US members of the Academy and assesses topics such as frequency of suits, demographies of defendants, disposition of suits, and plaintiff insurance status. The most recent survey took place in 2001 and found that 26% of pediatricians surveyed reported ever having a claim or suit brought against them (excluding claims during residency). Of those who had experienced a claim, 36% of the cases were settled out of court and 33% were dropped by the plaintiff. Only 3% of pediatricians reported that the plaintiff won their case.

A recent article published in Pediatrics by Kain and Caldwell-Andrews used data from the National Practitioner Data Bank (NPDB) to examine child-related malpractice payments. This study looked at malpractice payments involving children ≤19 years of age reported between February 1, 2004, and December 31, 2005, to the NPDB. Because the NPDB does not record data regarding the defendant’s specialty, the data reported include what can be considered high-risk specialties in terms of malpractice claims, such as obstetrics and gynecology. In fact 28% of the presented cases were obstetrics related. Inclusion of these types of claims most likely leads to inflated results in terms of the state of medical liability for pediatricians. Therefore, the Kain and Caldwell-Andrews article does not truly provide a clear picture of what impact malpractice is having on pediatricians.

The aim of this article was to examine malpractice claims data that are specific to the specialty of pediatrics, so as to provide a better understanding of the effect that the current malpractice crisis is having on this particular specialty and to determine whether this data provides any insight into how medical malpractice claims can be avoided.

METHODS
Selection of a Data Source
There are 3 primary sources of data available on medical malpractice judgments and awards: (1) NPDB, (2) Jury Verdict Research (JVR), and (3) Physician Insurers Association of America (PIAA) data-sharing project. Each of these data sources has its advantages and disadvantages.

The NPDB is maintained by the Department of Health and Human Services, and under the Health Care Quality Improvement Act of 1986, all payments in settlement of malpractice claims must be reported to this system. The NPDB is, therefore, the most comprehensive source of information about claims. Although data can be queried on the basis of plaintiff age, it is not possible to query the data on the basis of a specific physician specialty. Therefore, this data source cannot be used for this specific analysis, because physician specialty was our key search criterion.

The JVR collects data on jury verdicts only as reported to it by plaintiff’s attorneys, court clerks, and stringers. Because jury verdicts only represent a very small percentage of medical malpractice claims, and these verdicts are significantly higher than the average settlement, this data provides a skewed and limited picture of the medical malpractice system.
PIAA is a trade association of >50 medical malpractice insurance companies. Together, these companies insure ~60% of all private practicing physicians and surgeons in the United States. Since 1985, PIAA has maintained the PIAA data-sharing project, a database that pools information from >15 member-insurance companies, representing ~25% of the medical malpractice claims in the United States at a given time. All data are collected in a generic format where the identity of the individual claimants, insureds, and other parties are not recorded. These data are regularly analyzed for member organizations to document claim trends and identify underlying patient safety issues that result in malpractice cases. Although this database is not universally comprehensive, it does contain information not available in the NPDB, such as information on claims that are not ultimately paid and specialty of the defendant. Therefore, this data provides a broader view of the malpractice claims system than the NPDB.

Because the purpose of this article is to examine malpractice claims data that are specific to the specialty of pediatrics, the PIAA data-sharing project was chosen over the other 2 data sources.

Data Analysis
We asked PIAA to perform a query of its data-sharing project database looking at malpractice claims reported between January 1, 1985, and December 31, 2005, where the defendant’s medical specialty was coded as that of pediatrics. PIAA’s definition of pediatrics includes several different pediatric subspecialties, including neonatology, perinatology, and pediatric surgery. In addition, to compare pediatricians with other medical specialties, comparison data were pulled for the 27 other specialties recorded in the database using the same timeframe criteria.

Definitions
Terms are defined as follows:

- **Claim**: any written or oral demand for compensation in the form of money or services, with no legal paper having been filed in court. Policy provisions require insureds to notify the insurance company immediately on notice of a claim. Many claims that are unresolved later become suits.
- **Indemnity**: refers to settlements or awards made directly to plaintiffs as a result of claim-resolution process.
- **Medical misadventure**: descriptive terminology relating to an alleged principal departure from accepted medical practice. Twenty-eight misadventures are used in this study and can be broadly categorized as diagnosis related, procedure related, or case management related.
- **Severity**: severity of a patient’s injury is coded using the National Association of Insurance Commissioners severity codes 1 (emotional injury only) through 9 (death).

RESULTS
During a 20-year period (1985–2005), there were 214 226 closed claims reported to the PIAA data-sharing project. Pediatricians account for 2.97% of these claims (N = 6363), ranking it 10th among the 28 specialties in terms of the number of closed claims. Obstetric and gynecologic surgeons had the highest number of total closed claims at 29 453, accounting for 13.75% of the closed claims reported to PIAA. For pediatrics, 1790 of the closed claims were paid for a total indemnity of over $460 million. This makes pediatrics 16th in terms of indemnity payment rate (28.13%), with dentistry ranked highest at 43.35%, followed by obstetrics and gynecologic surgery at 35.50%. The indemnity payment rate for pediatrics has varied over time, with the lowest rate occurring in 2004 (20.14%) and the highest rate occurring in 1990 (33.45%) (Fig 1).

The mean indemnity payment for pediatrics between 1985 and 2005 was $261 263 with a median payment amount of $100 000, making it the seventh highest out of the 28 specialties. Neurosurgery had the highest median indemnity at $150 000 and oral surgery had the lowest at $13 500 (Fig 2). Indemnity payments have been steadily increasing over time for pediatrics. Using 2005 dollar values, the mean indemnity paid in 1985 was $232 987 with a median payment of $65 287. This has increased to an mean indemnity payment of $395 997 in 2005 and a median payment of $270 000 (Fig 3). Large loss claims (ie, those with an indemnity payment $500 000) made up 17.54% of paid claims. The majority of paid claims (54.47%) had an indemnity payment of less than $100 000.

Defense-related expenses were reported for 5402 of the 6363 closed pediatric claims between 1985 and 2005.
Total defense expenses for these claims came to more than $133 million with the mean expense for a pediatric claim being $24,634. Pediatrics ranked fourth highest among the 28 specialties in terms of mean expenses paid (Fig 4). Using 2005 dollar values, the mean expense for a malpractice claim has increased from a mean of $19,657 in 1985 to $40,851 in 2005 (Fig 5) with mean expenses in 2005 being considerably lower for cases where no indemnity was paid ($28,779) versus paid claims ($67,502). Between 1985 and 2005, the mean defense attorney expense for pediatric claims was $15,750, and the mean expert witness expense was $2,692. The ratio of expenses to indemnity for all closed cases between 1985 and 2005 was 28.45 cents expense per indemnity dollar.

Figure 6 shows the percentage of closed claims be-
between 1985 and 2005 by adjudication status for pediatrics. A total of 5.19% of closed claims resulted in a verdict, and the vast majority of these claims (79.01%) were won by the defense. The mean defense expense for claims that resulted in a defendant verdict was $59,147, with a median expense of $39,871. Of claims that resulted in a plaintiff verdict, the mean indemnity payment was $478,470 (median payment: $239,450), and mean defense expenses were $74,958 (median expense: $57,599). A total of 26.66% of claims resulted in a settlement in favor of the plaintiff. The mean indemnity in these cases was $250,650 (median indemnity: $100,000), and average defense-related expenses came to $33,710 (median expense: $20,973). Claims that were dropped, withdrawn, or dismissed had by far the lowest defense related expenses: $11,293 mean and $2747 median.

Most pediatricians involved with a claim reported to the PIAA data-sharing project between 1985 and 2005 were men (78.96%), board certified (81.51%), and had attended a US medical school (66.10%). A total of 60.9% had previous experience with a claim. A majority of claims involved care received either at a hospital location (49.01%) or in a practitioner’s office (43.41%).
Of those claims involving care received at a hospital, 69.77% of the hospitals were categorized as nonteaching institutions.

The 5 most common diagnoses involved in pediatric malpractice claims are presented in Table 1. It is important to note that these claims only represent 21.07% of closed claims reported between 1985 and 2005. Of the top 5, brain-damaged infant had the highest average indemnity payment and the largest average defense expenses payment. The severity of injury varied greatly when the medical diagnosis was routine infant or child health check, respiratory problems in the newborn, or appendicitis. However, 33.33% of meningitis related claims dealt with a plaintiff who had died, and in cases where the diagnosis was brain-damaged infant, the vast majority of plaintiffs (90.93%) had either major permanent injury or their condition was coded as grave.

The top 5 most prevalent medical misadventures for pediatrics are presented in Table 2. Errors in diagnosis are the most prevalent medical misadventure (31.87%) and include those claims involving a failure to diagnosis or an incorrect diagnosis. This type of misadventure also has the highest median indemnity payment and defense expenses. The most common medical condition associated with this medical misadventure was meningitis, but this only represents 8.17% of the claims coded as having an error in diagnosis as the medical misadventure.

The second most prevalent misadventure was no medical misadventure (21.63%), which refers to those claims that are believed to have legal merit but have no associated medical mishap. An example of this type of medical misadventure would be a claim that is filed where a newborn infant is found to have a brachial plexus injury that occurred during delivery. The claim might name several defendants, including the anesthesiologist who was in the room during the delivery, although there was nothing in the chart to show that the anesthesiologist was at fault. In this case, the anesthesiologist’s insurer would code this misadventure as no medical misadventure, because there was no evidence that the anesthesiologist or previous actions taken by the anesthesiologist had anything to do with the brachial plexus injury. This type of misadventure has the lowest indemnity payment rate of the top 5 most prevalent medical misadventures. The most common medical conditions associated with this medical misadventure in order of prevalence are brain-damaged infant (11.82%), routine infant or child health check (4.28%), meningitis (3.46%), respiratory problems in the newborn (2.56%), and birth (2.14%).

Improper performance, the third most common misadventure (13.25%), includes those claims where an act performed by a physician has resulted in a procedure being performed incorrectly. An example would be a circumcision procedure in which the foreskin was cut too short or the glans penis was injured resulting in deformity of the penis. The 2 most prevalent medical conditions associated with the medical misadventure of improper performance are brain-damaged infant (10.44%) and circumcision (5.11%).

A total of 9.46% of claims involve the failure of a physician to monitor the care of a patient or claims where the physician involved in a patient’s care has neglected the management of the treatment. These medical misadventures are labeled failure to supervise or monitor care, and they have the highest indemnity pay-

### Table 1: Five Most Prevalent Medical Conditions for Pediatrics (1985–2005)

<table>
<thead>
<tr>
<th>Medical Condition</th>
<th>Closed Claims, %</th>
<th>Indemnity Payment, $</th>
<th>Defense Expenses, $</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain-damaged infant</td>
<td>8.66</td>
<td>439,232</td>
<td>324,839</td>
</tr>
<tr>
<td>Meningitis</td>
<td>4.90</td>
<td>450,396</td>
<td>275,000</td>
</tr>
<tr>
<td>Routine infant or child health check</td>
<td>2.83</td>
<td>150,189</td>
<td>42,500</td>
</tr>
<tr>
<td>Respiratory problems in the newborn</td>
<td>2.56</td>
<td>270,607</td>
<td>162,500</td>
</tr>
<tr>
<td>Appendicitis</td>
<td>2.12</td>
<td>118,877</td>
<td>42,273</td>
</tr>
</tbody>
</table>

PEDIATRICS Volume 120, Number 1, July 2007
ment rate of the top 5 most prevalent medical misadventures. Once again, brain-damaged infant, respiratory problems in the newborn, and meningitis are among the most common medical conditions associated with this medical misadventure. However, asthma and convulsions are also among the top 5.

The fifth most prevalent medical misadventure was medication error (4.70%). This code applies to a claim if a medication problem is the central issue of a claim or if failure to prescribe is the central issue. The most prevalent medical conditions associated with this medical misadventure in decreasing order of prevalence are asthma, routine infant or child health check, heartburn, bronchitis, and convulsions.

In terms of total indemnity payments between 1985 and 2005, the 5 most expensive medical misadventures for pediatricians in decreasing order are errors in diagnosis, failure to supervise/monitor case, improper performance, no medical misadventure, and delay in performance. The medical misadventure of delay in performance refers to cases where the physician actually defers testing or treatment of a patient.

Table 3 presents data regarding the severity of a plaintiff’s injury between 1985 and 2005 for pediatrics. Most commonly, closed claims involve cases where death was the resultant injury (28.15%). Death also made up the largest percentage of paid claims (29.72%). However, grave injuries resulted in the largest percentage of total indemnity payments at 32.43% and a mean indemnity payment of $515,802.

**DISCUSSION**

Malpractice lawsuits are rightly the source of much consternation for physicians, and these reservations are often only worsened by anecdotal reports. In this article, we provide a great deal of data about the experience of pediatricians with malpractice and how pediatricians compare with other specialties. Although these data cannot remove the consternation physicians feel regarding the malpractice situation, it does serve to better inform the debate. Although pediatricians ranked 10th in paid claims reported, they represent <3% of paid claims. This number is somewhat less alarming than that reported by Kain and Caldwell-Andrews, who found that 14% of malpractice payments in 2004 involved a patient in the pediatric age group. We also found that in general, ~28% of cases result in an indemnity payment, at a median of $100,000. This data, in general, paints a much rosier picture than that seen in the Periodic Survey of Fellows in 2001.

Even more concerning to pediatricians than indemnity paid is what it might cost to defend themselves in a malpractice case. Although the costs of mounting a defense in cases should not be minimized, in 2005, they averaged less than $29,000 when no indemnity was paid and less than $68,000 when an indemnity was paid out. In addition, only ~5% of claims went to verdict, and only 20% of those claims (1% of claims overall) resulted in a verdict for the plaintiff.

The conditions most commonly seen in malpractice claims are brain-damaged infant and meningitis, but these represent only a very small percentage of all claims filed (~14%). The most common medical misadventure was error in diagnosis, followed by no medical misadventure. With no medical misadventure being this common, it is sure to cause many pediatricians to worry, because this means a significant number of cases involve

<table>
<thead>
<tr>
<th>Paid to Closed, %</th>
<th>Indemnity Payment, $</th>
<th>Defense Expenses, $</th>
</tr>
</thead>
<tbody>
<tr>
<td>Errors in diagnosis</td>
<td>31.87</td>
<td>260,015</td>
</tr>
<tr>
<td>No medical misadventure</td>
<td>21.63</td>
<td>332,442</td>
</tr>
<tr>
<td>Improper performance</td>
<td>13.25</td>
<td>204,737</td>
</tr>
<tr>
<td>Failure to supervise/monitor case</td>
<td>9.46</td>
<td>325,272</td>
</tr>
<tr>
<td>Medication errors</td>
<td>4.70</td>
<td>166,636</td>
</tr>
</tbody>
</table>

**TABLE 3** Patients’ Severity for Pediatric Malpractice Claims (1985–2005)

<table>
<thead>
<tr>
<th>Total Closed Claims, %</th>
<th>Total Paid Claims, %</th>
<th>Total Indemnity, %</th>
<th>Average Indemnity, $</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotional injury only</td>
<td>3.68</td>
<td>1.79</td>
<td>0.17</td>
</tr>
<tr>
<td>Insignificant injury</td>
<td>2.66</td>
<td>1.68</td>
<td>0.14</td>
</tr>
<tr>
<td>Minor temporary injury</td>
<td>10.56</td>
<td>8.04</td>
<td>2.42</td>
</tr>
<tr>
<td>Major temporary injury</td>
<td>10.18</td>
<td>8.77</td>
<td>5.23</td>
</tr>
<tr>
<td>Minor permanent injury</td>
<td>9.12</td>
<td>8.60</td>
<td>5.94</td>
</tr>
<tr>
<td>Significant permanent injury</td>
<td>10.45</td>
<td>12.29</td>
<td>13.8</td>
</tr>
<tr>
<td>Major permanent injury</td>
<td>12.45</td>
<td>12.68</td>
<td>21.26</td>
</tr>
<tr>
<td>Grave</td>
<td>12.76</td>
<td>16.42</td>
<td>32.43</td>
</tr>
<tr>
<td>Death</td>
<td>28.15</td>
<td>29.72</td>
<td>18.61</td>
</tr>
<tr>
<td>Totals</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>
no malpractice. It is important to note, however, that significantly less of these cases (only 6% of them) are actually paid. Although in an ideal world no cases without actual malpractice would result in payments or verdicts for the plaintiff, the good news is that these truly are a minority of cases in the malpractice system for pediatricians. However, it is also important to remember that all malpractice cases, even those that result in a favorable verdict for the physician, still have costs both emotionally and monetarily (ie, defense expenses).

As with all studies, there are limitations to this work that warrant consideration. The most obvious limitation lies with our choice of data source, the PIAA data-sharing project. This data source is not comprehensive, because it only represents ~25% of medical malpractice claims in the United States. However, the PIAA data-sharing project is the only option for analyzing claims by specialty and for examining claims that do not result in payment, which account for a very significant percentage of claims. It is also important to note that this study does not address the costs of malpractice insurance in any way. For many physicians, this is of even greater concern than malpractice lawsuits. Whether the increase in insurance premiums is warranted or appropriate is not the objective of this analysis and warrants additional study.

CONCLUSIONS
Malpractice is a serious issue, and one of concern for pediatricians and physicians. Many times, discussion of this issue is fueled only by anecdotes or generalizations. This study is not meant to end discussion on how to reform malpractice or improve care, and it is not meant to minimize the costs, both tangible and intangible, to physicians involved in malpractice cases. Instead, this study is meant to help inform the debate. There are some who will read the results of this analysis and draw comfort; others will view the same data with alarm and surprise. Regardless of how one interprets these findings, they are important in truly informing the debate with generalizable facts instead of politically motivated anecdotes.

ACKNOWLEDGMENTS
All data came from the PIAA data-sharing project (Rockville, MD). We thank Lori Bartholomew and Catherine Bernstein at PIAA for their assistance.

REFERENCES
Addressing Parents’ Concerns About Childhood Immunizations: A Tutorial for Primary Care Providers

Benjamin H. Levi, MD, PhD

Departments of Pediatrics and Humanities, Penn State College of Medicine, Hershey, Pennsylvania

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

ABSTRACT

BACKGROUND. Despite the dangers of vaccine-preventable infections and efforts by health care professionals to promote immunization, parents’ resistance to routine childhood immunizations continues to grow. This phenomenon can give rise to frustration among health care providers, as well as create barriers in providing medical care to children in need. In response, we developed a CD-ROM–based tutorial that (1) explains the nature and origins of parents’ concerns, (2) addresses clinical implications of resistance to immunization, (3) explores ethical and professional obligations that physicians have toward children and their parents, and (4) discusses how physicians can effectively address parents’ concerns.

OBJECTIVE. Our goals were to evaluate the tutorial’s effectiveness in improving physicians’ (1) general knowledge about parents’ resistance to childhood immunizations, (2) knowledge of adverse effects of immunization, and (3) attitudes toward parents’ resistance to childhood immunization.

DESIGN/METHODS. After pretesting, expert review, and revision, the 45-minute Penn State Immunization Project tutorial was pilot tested with pediatric and family medicine residents at 7 training programs in 4 states (Pennsylvania, New York, Maryland, and Iowa). Knowledge and attitudes were assessed by using a 26-item pretest/posttest, the results of which were then analyzed by using standard statistical methods.

RESULTS. A total of 122 residents completed the pretest/posttest. Statistically and clinically significant improvements were seen in residents’ general knowledge, knowledge of adverse events, and all 5 attitudinal measures regarding childhood immunizations.

CONCLUSIONS. The tutorial Addressing Parents Concerns About Childhood Immunizations: A Tutorial for Primary Care Providers is effective in improving resident physicians’ general knowledge, knowledge of adverse events, and attitudes. As such, this tutorial has the potential to enhance communication between parents and primary care providers and, more generally, improve clinicians’ response to the growing resistance toward routine childhood immunizations.
Despite the success of routine childhood immunizations, parental resistance continues to grow.1–13 Although often arising in geographic clusters,14,15 concern among parents is present in communities throughout the United States, as evidenced by news accounts, new legislation to eliminate state vaccination laws, and clinicians’ own experiences.16–18 This opposition is growing despite the dangers of vaccine-preventable infections,19–21 reassurance from researchers and oversight bodies as to the safety and efficacy of vaccines,22–26 and efforts by health care professionals to promote immunization.26–31 In part, this resistance is because of a proliferation of articles, books, and Web sites questioning the safety and value of routine childhood immunizations.10,34–37 Also contributing are the mixed messages parents receive from within the scientific and medical community,13,27,38–45 combined with parents’ confusion over the benefits versus risks of immunization.27,46–49

Another important ingredient, however, is the response that physicians provide when questioned by parents. Though pediatricians and family practice clinicians have a reputation as friendly and approachable, there are reports of parents having their concerns over immunization dismissed and/or disparaged, sometimes aggressively so.46 One recent study found that 24% to 39% of pediatricians reported they would dismiss a child from their practice if the parents refused ≥1 of the recommended vaccinations.50 It is not clear where these children would then go. But when parents’ concerns are not effectively addressed, often the end result is that children do not get the medical care they need and deserve.

To help clinicians better understand and deal with parents’ concerns over immunizations, we developed, through the Penn State Immunization Project (PSIP), a CD-ROM entitled Addressing Parents’ Concerns about Childhood Immunizations: A Tutorial for Primary Care Providers. This CD-ROM–based tutorial explains the nature and origins of parents’ concerns, including the historical, ideological, and scientific roots of resistance to immunization. It also addresses clinical implications of resistance to immunization, ethical and professional obligations physicians have toward children and their parents, and how physicians can effectively address parents’ concerns (see Table 1). The 2 primary goals of this tutorial are to increase knowledge and change attitudes about parental resistance to routine childhood immunizations.

The main target audience for the tutorial are primary care providers in their last stage of training (ie, pediatric and family practice residents, nurse practitioner students, etc), and community providers.

METHODS

Before developing the PSIP tutorial, we conducted a comprehensive review of (1) English-language professional and lay literature on opposition to childhood vaccinations (using PubMed, Google Scholar, and an extensive review of complementary and alternative medicine publications)33; (2) immunization handbooks and related government-sponsored educational materials from Australia, Britain, Canada, New Zealand, and the United States; and (3) online resources from organizations both in support of (eg, World Health Organization, National Network for Immunization Information, American Academy of Pediatrics) and opposition to (eg, National Vaccine Information Center) vaccination.

We identified allegations levied against routine childhood immunization practices, examined common responses to these challenges, and surveyed the relevant

---

TABLE 1 Overview of PSIP Tutorial

<table>
<thead>
<tr>
<th>1. Historical, ideological, and scientific roots of the issue</th>
</tr>
</thead>
<tbody>
<tr>
<td>i. Rationales</td>
</tr>
<tr>
<td>a. Ineffective</td>
</tr>
<tr>
<td>b. Injurious</td>
</tr>
<tr>
<td>c. Medical contraindications</td>
</tr>
<tr>
<td>d. Epidemiology</td>
</tr>
<tr>
<td>e. Unnatural</td>
</tr>
<tr>
<td>f. Religious objection</td>
</tr>
<tr>
<td>g. Political objection</td>
</tr>
<tr>
<td>h. Alternative health belief</td>
</tr>
<tr>
<td>ii. Research data</td>
</tr>
<tr>
<td>iii. Patient advocacy</td>
</tr>
<tr>
<td>iv. Ideological origins</td>
</tr>
<tr>
<td>v. Historical origins</td>
</tr>
<tr>
<td>vi. Social policy concerns</td>
</tr>
<tr>
<td>vii. Origins of parents’ concerns</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Clinical implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Research issues</td>
</tr>
<tr>
<td>B. Individual health concerns</td>
</tr>
<tr>
<td>C. Public health concerns</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Ethical and professional obligations</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Benefiting the patient</td>
</tr>
<tr>
<td>B. Curbing feelings of indignation</td>
</tr>
<tr>
<td>C. Shared decision-making</td>
</tr>
<tr>
<td>D. Policy concerns</td>
</tr>
<tr>
<td>i. Maximizing benefit</td>
</tr>
<tr>
<td>ii. Promoting justice</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. Responding to parents’ concerns</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Timeframe for working with parents</td>
</tr>
<tr>
<td>B. Resources: Professional and Public</td>
</tr>
<tr>
<td>C. Depicting/communicating risk</td>
</tr>
<tr>
<td>D. Professionalism</td>
</tr>
<tr>
<td>i. Identify points of disagreement/conflict</td>
</tr>
<tr>
<td>ii. Acknowledge limitations (personal, legal, relational)</td>
</tr>
<tr>
<td>iii. Trust (as a rate-limiting variable)</td>
</tr>
<tr>
<td>iv. Identify additional resources for parents</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>E. Putting concerns into context</th>
</tr>
</thead>
<tbody>
<tr>
<td>i. Evidence versus misinformation</td>
</tr>
<tr>
<td>ii. Credible versus non-credible sources</td>
</tr>
<tr>
<td>iii. Fear versus reasoned perspective</td>
</tr>
<tr>
<td>iv. Priorities</td>
</tr>
</tbody>
</table>

---
literature on risk communication and conflict management. We then used these themes to develop an alpha-version of the computer-based tutorial, and after obtaining institutional review board approval, tested it with a convenience sample of pediatric residents at the Penn State Hershey Medical Center. This testing included both a written pretest/posttest instrument (developed for this project) and a standardized patient exercise in which residents were videotaped (preintervention and postintervention) interacting with mothers (actually, paid actors) who resisted having their infants immunized. Based on test results and feedback from residents, we refined the alpha version of the tutorial and developed a beta version that was reviewed by experts in adult education, medical education, and professional development, as well as a group of 8 practicing pediatricians. Revisions were incorporated into the final 45-minute PSIP tutorial, which was then pilot tested with a convenience sample of pediatric and family medicine residents at 7 training programs in 4 states (Pennsylvania, New York, Maryland, and Iowa). The present version was delivered in a lecture format by the author; however, a freestanding version of the CD-ROM–based tutorial is in its final stage of development.

Residents’ knowledge and attitudes were assessed before and after the intervention using a written 26-item self-administered instrument. Twelve items assessed general knowledge about immunizations, such as immunization rates, impact of physician response on parental behavior, sources of opposition to vaccination, laws concerning immunization, and so on. Eight items assessed knowledge of adverse events from vaccination (proven versus alleged). All 20 of these items used a true/false/would-be-guessing scale. In addition to recording correct and incorrect responses, we also wanted to assess the magnitude of change, including responses of uncertainty (ie, would be guessing). As such, changes in preintervention versus postintervention responses were graded as shown in Table 2.

Five additional items used a 5-point Likert scale (strongly agree to strongly disagree) to measure residents’ attitudes regarding (1) the nature and etiology of parents’ resistance to vaccination and (2) the appropriateness of physician responses to such resistance. Each

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Pretest/Posttest Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td>Posttest</td>
</tr>
<tr>
<td>Incorrect</td>
<td>Incorrect</td>
</tr>
<tr>
<td>Incorrect</td>
<td>Guess</td>
</tr>
<tr>
<td>Incorrect</td>
<td>Correct</td>
</tr>
<tr>
<td>Guess</td>
<td>Guess</td>
</tr>
<tr>
<td>Guess</td>
<td>Incorrect</td>
</tr>
<tr>
<td>Guess</td>
<td>Correct</td>
</tr>
<tr>
<td>Correct</td>
<td>Correct</td>
</tr>
<tr>
<td>Correct</td>
<td>Incorrect</td>
</tr>
<tr>
<td>Correct</td>
<td>Guess</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 3</th>
<th>Demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
<td>n (%)</td>
</tr>
<tr>
<td>PGY 1</td>
<td>45 (37)</td>
</tr>
<tr>
<td>PGY 2</td>
<td>34 (28)</td>
</tr>
<tr>
<td>PGY ≥3</td>
<td>43 (35)</td>
</tr>
<tr>
<td>Pediatrics</td>
<td>60 (49)</td>
</tr>
<tr>
<td>Pediatrics/Internal medicine</td>
<td>12 (10)</td>
</tr>
<tr>
<td>Family practice</td>
<td>50 (41)</td>
</tr>
<tr>
<td>Iowa</td>
<td>29 (24)</td>
</tr>
<tr>
<td>Maryland</td>
<td>14 (11)</td>
</tr>
<tr>
<td>New York</td>
<td>42 (34)</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>37 (30)</td>
</tr>
</tbody>
</table>

PGY indicates postgraduate year. Due to rounding, numbers may not add to 100%.

increment of change along this scale toward the “desired” attitude was measured as +1, and each increment of change away from the desired attitude was measured as −1. Thus, the score for each item could range from −4 to +4. The mean change for each item was calculated by summing the positive and negative amounts of change, and averaging this sum by the number of individuals who changed their responses from the pretest to posttest. Because of how results were subsequently grouped, some changes (eg, from agree to strongly agree) may not be transparent in the aggregate report of pretest versus posttest results. The final test item identified the respondent’s residency discipline and year.

STATISTICAL ANALYSIS
Changes in preintervention versus postintervention responses were calculated as described in “Methods.” Descriptive statistics were generated including means, medians, and SDs for change variables and frequency tables for the original pretest and posttest responses. Differences between residency year/type and the outcome measures were characterized by using contingency table analysis; significance levels were determined by Pearson’s χ² statistic and 2 sample t tests.

RESULTS
In the 7 residency programs surveyed, there was a total of 281 residents. Of these, 122 (43%) were able to attend the lecture presentation, 100% of whom completed the pretest/posttest. Table 3 describes their distribution in terms of year and program.

Residents’ general knowledge improved significantly postintervention on 11 of 12 items (Table 4). Overall, 91% of postintervention responses were correct, compared with 50% of preintervention responses (see Fig 1). Fifty percent of the residents changed their response from pretest to posttest, with a mean change of +0.84 on a −1 to +1 scale. There were no statistically significant
Residents’ knowledge of adverse events also improved significantly on 5 of 8 items (Table 5). Overall, 89% of posttest responses were correct, compared with 56% pretest (see Fig 2). Forty-two percent of the residents changed their response from pretest to posttest, with a mean change of 0.72 on a 1 to +1 scale. There were no statistically significant differences based on respondents’ year in residency, type of residency program, or geographic location.

Change in attitude was assessed by measuring differences in preintervention/postintervention responses to the following statements, all 5 of which showed significant changes. For the statement “It is understandable...”

**TABLE 4** Change in General Knowledge

<table>
<thead>
<tr>
<th>Test Item (Correct Answer)</th>
<th>Percent Correct</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>In the United States, &gt;95% of 3- to 4-year-old children are up-to-date for their immunizations. (False)</td>
<td>57</td>
<td>85</td>
</tr>
<tr>
<td>Though primarily intended for use by physicians, the National Vaccine Information Center (NVIC) is also a good resource for parents. (False)</td>
<td>2</td>
<td>89</td>
</tr>
<tr>
<td>Opposition to vaccination is a fairly recent phenomenon, arising over the past 20–30 years. (False)</td>
<td>40</td>
<td>98</td>
</tr>
<tr>
<td>Evidence shows that physicians’ responses have little impact on whether parents vaccinate their children. (False)</td>
<td>64</td>
<td>98</td>
</tr>
<tr>
<td>For an injury to be reported and recorded in the Vaccine Adverse Events Reporting System, it must be shown to have been caused by a vaccine. (False)</td>
<td>31</td>
<td>79</td>
</tr>
<tr>
<td>Drug company funded research is much more likely (compared to independent research) to favor drug products being brought to market. (True)</td>
<td>67</td>
<td>96</td>
</tr>
<tr>
<td>Prominent physicians have voiced concern about the safety of routine childhood immunizations. (True)</td>
<td>41</td>
<td>98</td>
</tr>
<tr>
<td>In parts of the world, unsafe injection practices have linked immunization with hepatitis and HIV. (True)</td>
<td>44</td>
<td>92</td>
</tr>
<tr>
<td>In the United States, the risk of contracting vaccine preventable illnesses is still quite high. (False)</td>
<td>49</td>
<td>83</td>
</tr>
<tr>
<td>With regard to attending public school, state laws typically permit parents to refuse immunizations for their children if their decision is based on a religious belief. (True)</td>
<td>75</td>
<td>98</td>
</tr>
<tr>
<td>With regard to attending public school, state laws typically permit parents to refuse immunizations for their children if their decision is based on a philosophical (nonreligious) belief. (False)</td>
<td>37</td>
<td>80</td>
</tr>
<tr>
<td>State laws require that children receive all childhood immunizations recommended by the American Academy of Pediatrics. (False)</td>
<td>53</td>
<td>57</td>
</tr>
</tbody>
</table>

**TABLE 5** Change in Knowledge of Adverse Events

<table>
<thead>
<tr>
<th>There is good scientific evidence that in some children immunizations cause/precipitate... (Correct Answer)</th>
<th>Percent Correct</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anaphylaxis (True)</td>
<td>66</td>
<td>93</td>
</tr>
<tr>
<td>Inconsolable crying (up to 3–4 h) (True)</td>
<td>65</td>
<td>99</td>
</tr>
<tr>
<td>Sudden infant death syndrome (False)</td>
<td>72</td>
<td>95</td>
</tr>
<tr>
<td>Hypotonic reactions (incidence ~1 in 2000 kids) (True)</td>
<td>26</td>
<td>79</td>
</tr>
<tr>
<td>Anorexia (up to 48 h) (True)</td>
<td>33</td>
<td>91</td>
</tr>
<tr>
<td>Behavior disorders (chronic) (False)</td>
<td>74</td>
<td>79</td>
</tr>
<tr>
<td>Fever &gt;102°F (incidence 5%–10%) (False)</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>Seizures (incidence ~1%) (True)</td>
<td>23</td>
<td>23</td>
</tr>
</tbody>
</table>
that many parents are concerned about the risks of vaccination," 29% changed their response, 98% agreed or strongly agreed on the posttest, compared with 81% pretest, with a mean change of +1.31 (P < .0001; Fig 3). For the statement “Parents receive mixed messages about routine childhood immunizations,” 48% changed
their response, 99% agreed or strongly agreed on the posttest, compared with 78% pretest, with a mean change of +1.05 (P < .0001; Fig 4). For the statement “Most adverse reactions to routine childhood immunizations are insignificant by any measure,” 69% changed their response, 76% disagreed or strongly disagreed on the posttest, compared with 30% pretest, with a mean change of +1.67 (P < .0001; Fig 5). For the statement “Most parents who resist/oppose vaccinating their children are unreasonable,” 61% changed their response, 88% disagreed or strongly disagreed on the posttest, compared with 49% pretest, with a mean change of +1.28 (P < .0001; Fig 6). For the statement “Physicians are professionally justified in refusing to care for children whose parents oppose routine childhood immunizations,” 45% changed their response, 85% disagreed or strongly disagreed on the posttest, compared with 68% pretest, with a mean change of +1.07 (P < .0001; see Fig 7). There were no statistically significant differences based on respondents’ year in residency or type of residency program for any of these 5 attitudinal items. Geographic location, however, did matter in several state-to-state comparisons, with respondents from New York programs often showing less of a mean positive change. This was particularly true regarding the justifiability of refusing to provide care to those whose parents resisted immunization, with residents based in New York having a mean change of +0.06 compared with a mean change of +1.53 for residents from the other 3 states (P = .002).

DISCUSSION

Our educational tutorial was very successful in improving pediatric and family practice residents’ general knowledge and knowledge of adverse events regarding childhood immunizations. Although some residents were well informed about the historical and ideological origins of parents’ resistance to vaccination, most were not. The majority was unfamiliar with resources for accurate information, the nature of the reporting system for adverse events, physician involvement in the anti-vaccine movement, and the status of philosophical objections to immunization. A large portion of residents likewise had inaccurate beliefs about many of the clinical features that are (or are not) associated with routine childhood immunizations.

On all but 3 of the clinical measures, posttest knowledge improved significantly. For the 2 of 3 that did not (fever and seizures), it was subsequently noted that the wording and layout of the items in question lent themselves to misinterpretation. In the case of “Fever >102°F (incidence 5%–10%),” for example, 100% of respondents got this wrong on the posttest (the correct incidence being ~1%), presumably because the notation “(5%–10%)” had too small a font. That said, the combined knowledge posttest scores were significantly improved on 17 of 20 items.

Having a solid understanding of the clinical and contextual features that figure into parents’ resistance to immunization is crucial to respond appropriately to worried parents, the majority of whom simply are concerned about potential downsides of immunization and are confused about what and whom to believe.10,49,52–54 For those parents who resist/oppose immunization because of well-defined philosophical, religious, and/or alternative health beliefs, primary care providers must also understand the historical, ideological, and scientific roots of resistance to immunization.11–13,49,52,55 Our tutorial addresses these issues. It also addresses the clinical implications of resistance to immunization, how to communicate risk, how to deal with conflict that may arise,56–58 and how to identify reliable resources. Lastly, our tutorial discusses ethical and professional obligations with regard to children and parents, and how to effectively and respectfully help put parents’ concerns into context.59–61

Physicians are trained to be on the lookout for bad outcomes (including infectious disease) and to do what is feasible to prevent them. So, perhaps it is not too surprising that when a safe and effective tool such as immunizations is available, many physicians react with surprise if not consternation that someone would reject this offer of protection. The section of the tutorial on ethical and professional responsibilities is meant to address the tendency among some physicians to dismiss parents’ concerns as simply ignorant and/or confrontational.

Research has shown that how we choose to respond to parents’ concerns about immunizations significantly affects the course of action that parents choose, the partnership we have with them, and hence the quality of health care their children receive.15,57,60,62–65 A recent article examining parents’ experiences and decisions regarding immunizations provided the following depiction as not uncommon for parents who resist having their children immunized: “You just feel really painted into a corner and there’s really no support in the medical community. I went through...a dozen doctors who were just like, ‘I will not treat you if you’re not going to immunize your child’.”46 Nor are such professional reactions rare.50,66–67 Significant numbers of physicians hold that failure to immunize is tantamount to child abuse,66,67 and in 1 study the majority of county health directors believed that criminal charges and injunctions should be brought against parents for failing to immunize their child.59,61,68

For those who wish to better understand why it is a parent’s right (and responsibility) to decide what is in their child’s interests, including immunization, there is a growing literature that deals specifically with the ethics of immunization.61,69–74 Having already put into context how and why parents may have arrived at their resistance to immunization, the PSIP tutorial educates health
It is in this regard that the success of our tutorial is perhaps best demonstrated, namely, changes in attitude. After the tutorial, participants not only viewed parents’ resistance as more understandable and less unreasonable, given what parents had been exposed to, but also were more likely to believe that we are not professionally justified in refusing to care for children whose parents oppose routine childhood immunizations. On this final measure, 45% of participants reported some change in their attitudes toward parents who oppose immunizations, with only 15% maintaining that it is justified to refuse care because of opposition to immunization, compared with 32% before the tutorial. If the numbers from Flanagan-Klygis et al’s study\textsuperscript{50} are generalizable, namely, that 24% to 39% of physicians are willing to dismiss from their practice children whose parents refuse some or all immunizations, then our tutorial has the potential to help many physicians build more effective partnerships with the growing number of parents who resist immunizations.

Developing and nurturing a strong relationship with parents establishes trust, and without trust there is little we can do to help children and their parents. In a study that examined decision-making about vaccines, mothers who resisted immunization reported that despite seeking a trusting relationship with traditional pediatricians they felt alienated and turned away.\textsuperscript{46} As all pediatricians know, immunization is seldom an urgent matter. In the presence of a trusting relationship, we can normalize the issue as one of many we need to address, and reframe parents’ resistance in terms of our common interest in their child’s well-being. Our tutorial is designed to help physicians accomplish this goal.

We know from research that doctors who communicate effectively and respectfully are more likely to have their patients follow their recommendations and less likely to be sued.\textsuperscript{27,46,62,75–78} Our tutorial is not itself a tool for developing physicians’ communication skills, but it is intended to both provide tools to help them communicate with parents who resist immunization, and remind physicians that it is the job of parents to worry, and it is our job to help them figure out what to worry about.

LIMITATIONS
There are several limitations to this study. First, the calculations of mean change on the general knowledge and knowledge of adverse events items do not account for those who did not change their answer pretest to posttest. Second, the present study only examined knowledge and reported attitudes immediately before and after the intervention, not longitudinally. Thus, it is not clear how long the observed changes last. Third, because the study did not examine actual practice habits, it is not known how much behavior would change because of this intervention. In pilot testing of the alpha version of the tutorial, this was addressed with the use of videotaped interactions with “mock moms” preintervention and postintervention, including controls who had 2 preintervention encounters to control for improvement on repetition. Additional such testing would strengthen the present findings. Fourth, because of time and sample size constraints, the present study did not record demographic variables, and so cannot identify outcome differences between men and women, ethnic/racial groups, age cohorts, previous experience with parents, and so on. Fifth, the study included only residents, not board-certified physicians. Thus, it is not clear whether the current findings are generalizable to pediatricians and family medicine physicians out in practice. Additional studies are underway to examine this.

CONCLUSIONS
Routine childhood immunizations are among the most effective preventive health measures of the modern era. Despite their remarkable success, along with the very real dangers of vaccine preventable illnesses, there is growing concern among parents over the appropriateness of immunization. Our tutorial is an educational intervention designed to help physicians better understand the nature and origins of parents’ concerns, appreciate relevant ethical and professional issues, and develop a respectful and effective approach for working with parents who resist immunization. The tutorial is intended to be presented in person to physicians or other primary care providers and has sufficient support materials as to allow an interested individual to prepare himself or herself to give the presentation. Our data show that our tutorial is a very effective intervention for resident physicians. Additional study needs to be done to determine whether it is equally effective for physicians (and/or other primary care providers) already in practice.

ACKNOWLEDGMENTS
I thank Georgia Brown for invaluable assistance with this project, Sue Boehmer for conducting the statistical analysis, and the Robert Wood Johnson Foundation’s Generalist Faculty Physician Program for supporting the PSIP.

REFERENCES
1. Bellaby P. Has the UK government lost the battle over MMR? \textit{BMJ.} 2005;330:552–553
4. Kennedy AM, Brown CJ, Gust DA. Vaccine beliefs of parents


56. Kritek PB. Negotiating at an Uneven Table. 2nd ed. San Francisco, CA: Jossey-Bass; 2002


74. Veatch RM. The ethics of promoting herd immunity. Fam Community Health. 1987;10:44–53

75. Hickson GB, Clayton EW, Githens PB, Sloan FA. Factors that prompted families to file medical malpractice claims following perinatal injuries. JAMA. 1992;267:1359–1363


A Decline in the Frequency of Neonatal Exchange Transfusions and Its Effect on Exchange-Related Morbidity and Mortality

Laurie A. Steiner, MD, Matthew J. Bizzarro, MD, Richard A. Ehrenkranz, MD, Patrick G. Gallagher, MD

Division of Perinatal Medicine, Department of Pediatrics, Yale University School of Medicine, New Haven, Connecticut

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

ABSTRACT

OBJECTIVE. Our goal was to identify trends in patient demographics and indications for and complications related to neonatal exchange transfusion over a 21-year period in a single institution using a uniform protocol for performing the procedure.

METHODS. A retrospective chart review of 107 patients who underwent 141 single- or double-volume exchange transfusions from 1986–2006 was performed. Patients were stratified into 2 groups, 1986–1995 and 1996–2006, on the basis of changes in clinical practice influenced by American Academy of Pediatrics management guidelines for hyperbilirubinemia.

RESULTS. There was a marked decline in the frequency of exchange transfusions per 1000 newborn special care unit admissions over the 21-year study period. Patient demographics and indications for exchange transfusion were similar between groups. A significantly higher proportion of patients in the second time period received intravenous immunoglobulin before exchange transfusion. There was a higher proportion of patients in the 1996–2006 group with a serious underlying condition at the time of exchange transfusion. During that same time period, a lower proportion of patients experienced an adverse event related to the exchange transfusion. Although a similar percentage of patients in both groups experienced hypocalcemia and thrombocytopenia after exchange transfusion, patients treated from 1996–2006 were significantly more likely to receive calcium replacement or platelet transfusion. No deaths were related to exchange transfusion in either time period.

CONCLUSIONS. Improvements in prenatal and postnatal care have led to a sharp decline in the number of exchange transfusions performed. This decline has not led to an increase in complications despite relative inexperience with the procedure.
Exchange transfusion (ECT) was introduced in the late 1940s to decrease the mortality of hemolytic disease of the newborn (HDN) and to prevent kernicterus in surviving patients. ECT was subsequently applied to neonatal hyperbilirubinemia from a variety of causes and quickly became one of the most commonly performed neonatal procedures.

In 1968 and 1971, Lucey accurately predicted that prenatal interventions, particularly the development of Rh-immunoglobulin, coupled with advances in postnatal care such as phototherapy, would lead to a dramatic decline in the number of ECTs performed. Maisels, in a review that combined data from 3 centers over 40 years, observed a decline in the frequency of ECT and predicted that it would lead to increased complications because of inexperience with the procedure. More recent advances, such as use of intratuterine transfusions and improvements in diagnostic ultrasound, have likely accelerated such as use of intrauterine transfusions and improvement in the frequency of ECT.

Since the introduction of ECT, the level of bilirubin at which to initiate this procedure has been a controversial issue. Based on experience with HDN, a bilirubin level of 20 mg/dL was used by many centers, including Yale, but some questioned whether it was appropriate to apply this cutoff to patients with nonhemolytic hyperbilirubinemia. This debate intensified in the late 1980s when several reports demonstrated that term infants with nonhemolytic jaundice were not as susceptible to kernicterus as infants with HDN.

In 1994, the American Academy of Pediatrics (AAP) published its first guidelines on the treatment of hyperbilirubinemia. These guidelines increased the bilirubin threshold for initiating ECT in term infants without hemolysis and allowed for a trial of intensive phototherapy before an ECT was initiated. In addition, these guidelines encouraged prenatal testing of maternal ABO and Rh types and recommended increased monitoring for hyperbilirubinemia in all infants. These interventions had the potential to cause a further decline in the number of patients requiring ECT.

We hypothesized that changes in prenatal and postnatal care have altered the patient population undergoing ECT, the indication for exchange, and the incidence of ECT-related morbidity and mortality. To examine this, we performed a longitudinal, 21-year review of ECT at a single center.

**Patients and Methods**

Infants who required single- or double-volume ECT and had long-term admissions (>24 hours) in the newborn special care unit (NBSCU) at Yale New Haven Hospital (YNHH) from January 1, 1986, through December 31, 2006, were included. Neonates who received partial ECT for polycythemia or anemia were excluded. Data collection included patient demographics, comorbidities, indication for exchange transfusion, treatment with phototherapy and intravenous immunoglobulin (IVIg), and ECT-related complications.

Patients were divided into 2 groups, 1986–1995 and 1996–2006, based on the AAP guidelines for the management of hyperbilirubinemia published in October 1994 and implemented in the NBSCU at Yale in late 1995. Before this time, the threshold for ECT at YNHH, with or without evidence of hemolysis, was a total serum bilirubin of 20 mg/dL for term infants, with threshold levels decreasing based on birth weight. Beginning in late 1995 and continuing to the present time, the threshold for ECT at YNHH was raised to 25 mg/dL for term infants >48 hours old without evidence of hemolysis, but remained at 20 mg/dL for those with hemolysis. Asymptomatic term infants were also provided the opportunity to respond to intensive phototherapy before an ECT was initiated, and all infants were strictly monitored for hyperbilirubinemia as per the AAP guidelines.

A detailed, step-by-step protocol, provided in the YNHH NBSCU procedure manual, was used for ECT. This technique, as described by Edwards and Fletcher, did not change over the 21-year study period.

**Indications and Comorbidities**

The indications for ECT were hyperbilirubinemia or anemia. Hyperbilirubinemia was further classified by etiology (Rh disease, ABO incompatibility, idiopathic hyperbilirubinemia, and other hematologic diagnoses). Patients were considered to have a significant preexisting comorbidity if they were treated with blood pressure support and/or mechanical ventilation, if they had a major congenital anomaly, or if they had any of the following diagnoses: respiratory distress syndrome, intraventricular hemorrhage (all grades as defined by Papile et al), necrotizing enterocolitis (NEC; modified Bell’s criteria at least stage 2a), or sepsis (defined as a positive blood culture and/or signs and symptoms consistent with sepsis treated with antibiotics for ≥7 days).

**ECT-Related Complications**

ECT-related complications were defined as any complication, not present before the ECT, which occurred within 7 days after the exchange. They were defined as follows: severe thrombocytopenia, platelet count <50 000/mm³; hypocalcemia, serum calcium <8.0 mg/dL or plasma ionized calcium <3.5 mg/dL; seizures, clinical evidence of seizure-like activity treated with antiseizure medication; bradycardia, heart rate <100 beats per minute; apnea, cessation of respirations for >20 seconds; catheter malfunction, central venous or arterial catheter thrombosis or rupture; hyperkalemia, serum potassium >6.5 meq/dL associated with electrocardiogram changes; NEC, modified Bell’s criteria at least stage 2a diagnosed after the ECT; and ECT-related mortality, ECT-related mortality was defined as any death that was
directly related to the ECT and occurred within 7 days after the exchange.

**Statistical Analysis**

SPSS 13.0 (SPSS Inc, Chicago, IL) and GraphPad Prism 3.0 (GraphPad Software, Inc, San Diego, CA) were used for data analyses. Continuous data were compared by using the Student’s *t* comparison of means. Dichotomous data were compared by using a Pearson’s χ² analysis or Fisher’s exact test when at least 1 cell contained a value <5. Trends were analyzed by using linear regression analysis. To incorporate both inborn and outborn neonates into this analysis of trends, the number of ECTs was evaluated per 1000 NBSCU admissions. In evaluating inborn neonates separately, the number of ECTs was evaluated per 1000 live births. A *P* value of <.05 was considered statistically significant.

This study was approved by the institutional review board of the Yale University School of Medicine.

**RESULTS**

From January 1, 1986, to December 31, 2006, there were 98,901 live births at YNHH and 16,389 long-term admissions, inborn and outborn, to the NBSCU. One hundred seven infants underwent 141 ECTs from 1986–2006. Two patients in each time period received a single-volume ECT, with the remaining patients receiving a double-volume or near–double-volume exchange. Over the entire study period, there was a statistically significant decline in the number of ECTs performed per 1000 live births in inborn neonates (*r*² = 0.30; *P* = .010; Fig 1A) and per 1000 NBSCU admission in those both inborn and outborn (*r*² = 0.49; *P* < .001; Fig 1B).

Demographic data were similar between the 2 groups, with no statistically significant differences in gestational age, birth weight, race, gender, or age at ECT (Table 1). The rate of phototherapy before exchange did not differ significantly between groups. Neonates in the 1996–2006 group were significantly more likely to receive IVIg before ECT (*P* = .016; Table 1).

There were no statistically significant differences in the indications for ECT when comparing the 1986–1995 and 1996–2006 groups (Table 2). The most common indication for ECT was hyperbilirubinemia, which was further subdivided into ABO incompatibility, Rh disease, idiopathic hyperbilirubinemia, and other hematologic diagnoses. Other diagnoses included glucose-6-phosphate dehydrogenase deficiency, pyruvate kinase deficiency, fibrosarcoma with large vascular compartment, hemolytic anemia because of Gram-negative sepsis, congenital acute myelogenous leukemia, α-thalassemia, hereditary pyropoikilocytosis, and hereditary spherocytosis. The most common cause of hyperbilirubinemia requiring ECT was Rh disease. Antibodies to non-D Rh antigens were common, occurring in 40% of patients with Rh disease in the first group and in 64% of patients with Rh disease in the second group.

A smaller proportion of patients in the 1996–2006 group experienced an ECT-related complication (Table 3). This result was not statistically significant, possibly because of the small sample size. We observed a high rate of thrombocytopenia and hypocalcemia after ECT in both the 1986–1995 and the 1996–2006 groups, comparable to previous studies. Despite similar rates of thrombocytopenia and hypocalcemia, patients treated from 1996–2006 were significantly more likely to be transfused platelets or to be given intravenous calcium (Table 4). The retrospective nature of this study and the small sample size make it difficult to determine the causality of these observations. The higher proportion of preexisting comorbidities in the neonates undergoing ECT from 1996 to 2006 may have resulted in more aggressive management or, alternatively, the difference might stem from unidentified changes in our clinical practice over the last 2 decades.

A total of 5 deaths occurred within 7 days of the ECT, none of which were related to the ECT.
Authors of previous reports have hypothesized that premature infants are more susceptible to complications from ECT. We observed no significant differences in either time period in the frequency of ECT-related complications or their treatment in neonates with birth weight compared with those <1500 g. Infants <1500 g did not experience increased rates of thrombocytopenia, hypocalcemia, calcium replacement, or platelet transfusion (data not shown). The small sample size of this premature cohort made it difficult to draw any valid conclusions from the analyses.

**DISCUSSION**

These data demonstrate a dramatic decline in the frequency of ECT at YNHH over 2 decades, representing the longest single-center, longitudinal documentation of trends in ECT. This decline is likely multifactorial with contributions from advances in both prenatal and postnatal care, such as middle cerebral artery Doppler studies to noninvasively follow fetal anemia, and IVIg treatment for patients with hemolysis. In addition, adoption of the 1994 AAP guidelines may have contributed to this decline.

---

**TABLE 1**

Demographic Data and Age at Exchange Transfusion

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gestational age, mean ± SD, wk</td>
<td>35.3 ± 4.7</td>
<td>35.7 ± 4.8</td>
<td>34.6 ± 4.5</td>
<td>.257</td>
</tr>
<tr>
<td>Birth weight, mean ± SD, g</td>
<td>2511.1 ± 983.8</td>
<td>2469.4 ± 956.4</td>
<td>2593.1 ± 1044.7</td>
<td>.541</td>
</tr>
<tr>
<td>Birth weight &lt;1000 g, n (%)</td>
<td>12 (11)</td>
<td>7 (10)</td>
<td>5 (14)</td>
<td>.747</td>
</tr>
<tr>
<td>Birth weight &lt;1500 g, n (%)</td>
<td>19 (18)</td>
<td>12 (17)</td>
<td>7 (19)</td>
<td>.740</td>
</tr>
<tr>
<td>Male gender, n (%)</td>
<td>61 (57)</td>
<td>43 (61)</td>
<td>18 (50)</td>
<td>.296</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>63 (59)</td>
<td>43 (61)</td>
<td>20 (56)</td>
<td>.617</td>
</tr>
<tr>
<td>Black</td>
<td>31 (29)</td>
<td>20 (28)</td>
<td>11 (31)</td>
<td>.791</td>
</tr>
<tr>
<td>Hispanic</td>
<td>9 (8)</td>
<td>5 (7)</td>
<td>4 (11)</td>
<td>.715</td>
</tr>
<tr>
<td>Asian</td>
<td>4 (4)</td>
<td>3 (4)</td>
<td>1 (3)</td>
<td>.999</td>
</tr>
<tr>
<td>Transport, n (%)</td>
<td>34 (32)</td>
<td>23 (32)</td>
<td>11 (31)</td>
<td>.841</td>
</tr>
<tr>
<td>Age at exchange, mean ± SD, d</td>
<td>3.6 ± 3.1</td>
<td>3.4 ± 2.9</td>
<td>4.0 ± 3.6</td>
<td>.347</td>
</tr>
<tr>
<td>Phototherapy before ECT, n (%)</td>
<td>91 (85)</td>
<td>60 (85)</td>
<td>31 (89)</td>
<td>.823</td>
</tr>
<tr>
<td>IVIg administration, n (%)</td>
<td>6 (17)</td>
<td>1 (1)</td>
<td>5 (14)</td>
<td>.016</td>
</tr>
<tr>
<td>Intrauterine transfusions, n (%)</td>
<td>20 (19)</td>
<td>14 (20)</td>
<td>6 (17)</td>
<td>.699</td>
</tr>
<tr>
<td>Comorbidities, n (%)</td>
<td>41 (38)</td>
<td>24 (34)</td>
<td>17 (47)</td>
<td>.177</td>
</tr>
<tr>
<td>Mechanical ventilation</td>
<td>37 (35)</td>
<td>23 (32)</td>
<td>14 (39)</td>
<td>.502</td>
</tr>
<tr>
<td>Blood pressure support</td>
<td>17 (16)</td>
<td>11 (15)</td>
<td>6 (17)</td>
<td>.888</td>
</tr>
<tr>
<td>NEC</td>
<td>4 (4)</td>
<td>1 (1)</td>
<td>3 (8)</td>
<td>.110</td>
</tr>
<tr>
<td>Hydrops fetalis</td>
<td>6 (6)</td>
<td>3 (4)</td>
<td>3 (8)</td>
<td>.661</td>
</tr>
<tr>
<td>IVH</td>
<td>10 (9)</td>
<td>8 (11)</td>
<td>2 (6)</td>
<td>.490</td>
</tr>
<tr>
<td>RDS</td>
<td>23 (21)</td>
<td>14 (20)</td>
<td>9 (25)</td>
<td>.532</td>
</tr>
<tr>
<td>Sepsis</td>
<td>13 (12)</td>
<td>7 (10)</td>
<td>6 (17)</td>
<td>.354</td>
</tr>
</tbody>
</table>

* Comparison of the 2 time periods.

**TABLE 2**

Indication for ECT

<table>
<thead>
<tr>
<th></th>
<th>Total (N = 141), n (%)</th>
<th>1986–1995 (N = 96), n (%)</th>
<th>1996–2006 (N = 45), n (%)</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperbilirubinemia</td>
<td>120 (85)</td>
<td>79 (82)</td>
<td>41 (91)</td>
<td>.211</td>
</tr>
<tr>
<td>Rh disease</td>
<td>58 (41)</td>
<td>41 (43)</td>
<td>17 (39)</td>
<td>.578</td>
</tr>
<tr>
<td>ABO incompatibility</td>
<td>39 (28)</td>
<td>28 (29)</td>
<td>11 (24)</td>
<td>.560</td>
</tr>
<tr>
<td>Idiopathic</td>
<td>28 (20)</td>
<td>17 (18)</td>
<td>11 (24)</td>
<td>.351</td>
</tr>
<tr>
<td>Other</td>
<td>14 (10)</td>
<td>10 (10)</td>
<td>4 (9)</td>
<td>.999</td>
</tr>
<tr>
<td>Anemia</td>
<td>3 (2)</td>
<td>1 (1)</td>
<td>2 (4)</td>
<td>.239</td>
</tr>
</tbody>
</table>

* Comparison of the 2 time periods.

**TABLE 3**

Exchange Transfusion-Related Complications Excluding Thrombocytopenia and Hypocalcemia

<table>
<thead>
<tr>
<th></th>
<th>Total (N = 141), n (%)</th>
<th>1986–1995 (N = 96), n (%)</th>
<th>1996–2006 (N = 45), n (%)</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catheter malfunction</td>
<td>4 (3)</td>
<td>2 (2)</td>
<td>2 (4)</td>
<td>.592</td>
</tr>
<tr>
<td>Seizures</td>
<td>3 (2)</td>
<td>3 (3)</td>
<td>0 (0)</td>
<td>.551</td>
</tr>
<tr>
<td>NEC</td>
<td>2 (1)</td>
<td>2 (2)</td>
<td>0 (0)</td>
<td>.562</td>
</tr>
<tr>
<td>Anemia</td>
<td>1 (1)</td>
<td>1 (1)</td>
<td>0 (0)</td>
<td>.999</td>
</tr>
<tr>
<td>Bradycardia</td>
<td>5 (4)</td>
<td>4 (4)</td>
<td>1 (2)</td>
<td>.673</td>
</tr>
<tr>
<td>Hyperkalemia</td>
<td>1 (1)</td>
<td>1 (1)</td>
<td>0 (0)</td>
<td>.999</td>
</tr>
<tr>
<td>Any complication</td>
<td>16 (11)</td>
<td>13 (14)</td>
<td>3 (7)</td>
<td>.270</td>
</tr>
</tbody>
</table>

* Comparison of the 2 time periods.

**TABLE 4**

Hypocalcemia and Thrombocytopenia in Patients Undergoing Exchange Transfusion

<table>
<thead>
<tr>
<th></th>
<th>Total (N = 141), n (%)</th>
<th>1986–1995 (N = 96), n (%)</th>
<th>1996–2006 (N = 45), n (%)</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypocalcemia</td>
<td>53 (38)</td>
<td>32 (33)</td>
<td>21 (47)</td>
<td>.128</td>
</tr>
<tr>
<td>Calcium replacement</td>
<td>24 (45)b</td>
<td>9 (28)</td>
<td>15 (71)</td>
<td>.002</td>
</tr>
<tr>
<td>Thrombocytopenia</td>
<td>53 (38)</td>
<td>36 (38)</td>
<td>17 (38)</td>
<td>.999</td>
</tr>
<tr>
<td>Platelet transfusion</td>
<td>26 (49)c</td>
<td>11 (31)</td>
<td>15 (88)</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

* Comparison of the 2 time periods.
b Percentage with hypocalcemia who received calcium replacement.
c Percentage with thrombocytopenia who received platelet transfusion.
The 1994 AAP guidelines recommend that all infants jaundiced in the first 24 hours of life receive a total serum bilirubin and all infants be assessed for jaundice by a health care provider at 2 to 3 days of life. These guidelines also recommend prenatal testing of maternal ABO and Rh types, prenatal screening for unusual maternal antibodies, and screening of the cord blood if the mother was Rh negative or if the mother’s ABO type was unknown. The heightened monitoring of all infants for hyperbilirubinemia may have contributed to early detection and treatment of infants with significant jaundice (hemolytic and nonhemolytic) and, therefore, caused a decline in the number of ECT necessary.

The declining rate of ECT has led to speculation that inexperience with the procedure would result in increased rates of ECT-associated morbidity and mortality. Historically, morbidity and mortality associated with ECT steadily declined from the 1950s through the 1970s, a period of time when the procedure was commonly performed, seeming to reach a nadir in the mid-1970s when Keenan et al reported a serious adverse event rate of 5.2% and a mortality rate of 0.5%. As ECT became a less common occurrence in the NICU, reports of ECT-related morbidity and mortality were also less frequent, but described higher rates of ECT-related complications and mortality. In 1997, Jackson reported data on 106 neonates from 2 NICUs who underwent 140 ECTs from 1980–1995. He described a high rate of serious complications (12%) in the 25 neonates with significant preexisting comorbidities, as well as an overall mortality attributable to ECT of 2%. In 2004, Patra et al reported data on 55 neonates from 2 high-volume NICUs who underwent 66 ECTs from 1992–2002. Although their rate of ECT was quite low (<3 ECTs in each unit per year), they reported a high rate of ECT-related adverse events (74%) and a mortality rate of 2%.

In contrast to these smaller studies, our overall rate of ECT-related complications did not increase throughout the study period despite a sharp decline in the number of ECTs performed. The majority of complications were transient and, similar to previous studies, when serious adverse events occurred, they were observed in patients with significant, preexisting comorbidities. In addition, no cases of ECT-related mortality were observed. We speculate that a combination of factors likely contributed to this observation, including the use of a standardized ECT protocol at our institution and increasing attending neonatologist involvement in ECT procedures. The impression of 1 of the authors (Dr Ehrenkranz), who worked in the NBSCU during the entire 21-year study period, is that attending supervision has increased as the number of ECTs declined and has become nearly universal.

In an era when many neonatal care givers have more experience with advanced therapies such as high-frequency ventilation, dialysis, and extracorporeal membrane oxygenation than with ECT, a standardized protocol for performing ECT may be an important tool for decreasing the number of adverse, procedure-related events. Inclusion of ECT in neonatal education will also help minimize ECT-related morbidity and mortality, even as the frequency of ECT continues to decline.

ACKNOWLEDGMENT
This work was supported, in part, by National Institute of Child Health and Human Development grant T32 HD07094 (to Dr Steiner).

REFERENCES
READ THIS AND WEEP: CRYING AT WORK GAINS ACCEPTANCE

“Crying at work has long been seen as verboten. But there’s evidence that a growing number of workers, especially those in their 20s and 30s, see it differently. Some think it’s old-fashioned to hide our emotions. Others are quick to cry over negative feedback. And many find themselves at odds with managers who grew up with a more repressive definition of professional conduct. . . . Savvy bosses also avoid jumping to the conclusion that tears signal weakness. In a survey of 182 medical students several years ago, Nancy Angoff, an associate dean at the Yale School of Medicine, found 133 had cried at least once during clinical training, for reasons ranging from stress or mistreatment to compassion and empathy for patients. Instructors ‘need to acknowledge that it is not only OK to cry,’ she wrote, ‘but it is understandable, appropriate and sometimes desirable.’”


Noted by JFL, MD
Characteristics of the Pediatric Hospitalist Workforce: Its Roles and Work Environment

Gary L. Freed, MD, MPH\textsuperscript{a,b}, Kathryn Brzoznowski, MPH\textsuperscript{a,b}, Kamilah Neighbors, MHSA\textsuperscript{a,b}, Indu Lakhani, MBA, MS\textsuperscript{a,b}, American Board of Pediatrics, Research Advisory Committee

\textsuperscript{a}Child Health Evaluation and Research Unit, \textsuperscript{b}Division of General Pediatrics, University of Michigan, Ann Arbor, Michigan

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

\section*{ABSTRACT}

\textbf{OBJECTIVE.} Over the past 10 years, the use of hospitalists has grown in both the adult and pediatric setting as a response to pressure to deliver cost-effective, high-quality care. However, there is a paucity of information regarding the variation in the clinical roles, educational responsibilities, work patterns, and employment characteristics of pediatric hospitalists. This lack of information hampers efforts to define the nature of the field and determine whether any formalized, additional training or experience should be required for physicians in this clinical practice domain.

\textbf{DESIGN.} We conducted a telephone survey of a national sample of pediatric hospitalist program directors ($n = 116$). Questionnaire items focused on exploring the clinical roles, work patterns, employment characteristics, and training of pediatric hospitalists within each institution. Results were stratified by teaching hospitals, urban/rural location, hospital size, and membership in the National Association of Children’s Hospitals and Related Institutions.

\textbf{RESULTS.} The response rate was 97%. The majority of hospitals surveyed (70\%) reported that hospitalists do not generate enough income from professional billing to pay their salaries. Fewer than half (39\%) of respondents reported that their hospital measures pediatric clinical outcomes associated with hospitalist care. A total of 42\% of hospitalist program directors reported that most of their hospitalists had an average duration of employment of $<3$ years. In programs with residents, hospitalists serve as teaching attendings for pediatric patients in almost all cases (89\%).

\textbf{CONCLUSIONS.} Hospital medicine is a rapidly growing enterprise. A better understanding of both its participants, as well as those affected by its practice, will enable planning for a future that meets as many needs as possible while ensuring the best possible care for children.
The Society of Hospital Medicine defines a hospitalist as a physician whose primary focus is the general medical care of hospitalized patients and whose responsibilities also include teaching, research, and leadership related to hospital care. Over the past 10 years, the use of hospitalists has grown in both the adult and pediatric setting as a response to pressure to deliver cost-effective, high-quality care. Whether the training and experience of the physicians placed in such roles is commensurate with those expectations remains unknown.

In 2002, there were ~600 pediatric hospitalists, with many more believed to be active today. However, there is a paucity of information regarding the variation in the clinical roles, educational responsibilities, work patterns, and employment characteristics of pediatric hospitalists. This lack of information hampers efforts to define the nature of the field and to determine whether any formalized, additional training or experience should be required for physicians in this clinical practice domain.

We conducted this national study of hospitalist program directors to better understand the current nature of the field and the degree of variation present among hospitalists and the programs in which they are employed.

**METHODS**

**Study Sample**

From the American Hospital Association’s 2005 annual survey of hospitals, the research team identified 761 hospitals that reported having both a hospitalist service and pediatric beds. From these 761 hospitals, the research team sought a sample of 200, stratified by the following variables:

- Council of Teaching Hospitals (COTH) designation (teaching versus nonteaching);
- National Association of Children’s Hospitals and Related Institutions (NACHRI) designation (children’s hospital versus non–children’s hospital);
- freestanding (freestanding versus part of hospital system);
- metropolitan statistical area size (urban versus rural); and
- hospital size (small [<250 total beds] versus large [≥250 total beds]).

The final sample was 213. We selected an additional 13 hospitals to ensure appropriate sampling across stratification variables.

Hospitals were sampled with varying probabilities from each stratum. Weights were applied to create a representative sample of the overall hospital population. The total sampling weight (TSW) calculated for each hospital was based on the probability of selection into the study ($P$) and the response rate (RR). We used the formula $\text{TSW} = (1/P) \times (1/RR)$.

**Survey Instrument**

We constructed a 42-item structured questionnaire to be administered by telephone. The instrument was designed to be completed in 15 minutes or less. Questionnaire items explored the clinical roles, work patterns, employment characteristics, and training of pediatric hospitalists within each institution.

**Questionnaire Administration**

From June through September 2006, members of the research team telephoned the sampled hospitals to determine whether the hospital had pediatric hospitalists, and if so, to identify the physician responsible for each program. Most individuals were either chiefs or chairs of pediatric departments, or the head or most senior pediatric hospitalist at the institution. A study introduction letter was sent to those individuals. Research staff then contacted potential respondents, explained the purpose of the study, and obtained verbal consent to begin the survey.

**Data Analysis**

Initially, unweighted frequency distributions were calculated for all survey items as descriptive statistics. Next, weighted $\chi^2$ statistics were used to determine levels of association between the survey responses and the various hospital classifications listed above. For the comparisons between classifications of hospitals, only weighted percentages are provided. The study was approved by the University of Michigan Medical institutional review board.

**RESULTS**

**Response Rate**

Eligible respondents had to have ≥1 pediatric hospitalist, defined as a physician who is not running a subspecialty service and whose primary professional focus is the general medical care of hospitalized children. Ninety-seven hospitals were not eligible because they reported having no pediatric hospitalist, despite having been identified in the American Heart Association database as having both a hospitalist service and pediatric beds. These hospitals only had hospitalists that cared for adult patients. Of the remaining eligible hospitals ($n = 116$), 112 completed the telephone survey, representing an overall response rate of 97%. Characteristics of participating and noneligible hospitals from our original sample are found in Table 1.

**Hospital Demographics**

Among the 112 hospitals with completed telephone surveys, 48% ($n = 54$) were designated as teaching hospi-
The use of pediatric hospitalists at most hospitals is a recent trend; a majority (55%) have been using pediatric hospitalists for ≤5 years. (Table 2).

Fewer than half (44%; n = 49) of those surveyed reported that their hospital had a formal definition of a hospitalist. Of these 49 hospitals, the most common definitions given were consistent with the Society of Hospital Medicine1 (39%; n = 18), followed by a more restrictive definition of only providing inpatient care (37%; n = 17); 9% (n = 4) defined a hospitalist as a physician that spends ≥25% of their clinical time providing hospital-based care, whereas the remaining 15% (n = 7) did not know their hospital’s definition.

The majority of hospitals surveyed had small pediatric hospitalist programs; 59% (n = 66) had 1 to 5 physicians serving this role (Table 3); 49% plan to increase the number of full-time employees of pediatric hospitalists in the next year.

The proportion of pediatric inpatients cared for by hospitalists varies widely across institutions. Only 4% (n = 5) of hospitals report that hospitalists care for 100% of their pediatric inpatient population, whereas 37% (n = 41) report hospitalists care for less than one quarter of their pediatric inpatients. There were no significant differences across hospital types.

The majority of pediatric hospitalists are employed full-time: 63% (n = 70) of programs report that at least three quarters of their pediatric hospitalists are employed full-time as hospitalists in clinical care. Only 16% of programs had no full-time hospitalists. There were no differences across hospital types.

In the majority (64%; n = 72) of programs, full-time hospitalists work >40 hours per week at the hospital on-site when on service, and in 34% (n = 38) they take call from home 1 to 2 nights per week. Nearly half (36%; n = 40) of hospitals report their hospitalists spend ≤26 weeks on service each year.

### Clinical Roles

Nearly all respondents (96%; n = 107) reported that hospitalists provide service in the general pediatric inpatient unit during the day. Another common service covered by pediatric hospitalists was inpatient general medical consultation (88%; n = 98). Fewer hospitals reported pediatric hospitalist coverage in each of the following clinical settings: subspecialty pediatric inpatient service (38%; n = 43); normal newborn nursery (38%; n = 43); emergency department (37%; n = 41); outpatient or outreach clinics (20%; n = 22); PICU (18%; n = 20); NICU (11%; n = 12); and transports (8%; n = 9).

Regarding procedures performed or supervised by pediatric hospitalists, lumbar punctures (96%; n = 107) and sedation services (62%; n = 69) were common, whereas infusion services (26%; n = 29), percutaneous intravenous catheter line placement (26%; n = 29), and circumcision (16%; n = 18) were performed less commonly.

The number of medical and surgical beds covered was distributed fairly evenly; approximately one third of hospitals reported that hospitalists covered <15 beds (30%; n = 34), one third reported that 15 to 30 beds were covered (34%; n = 38), and 31% (n = 35), reported coverage of ≥30 beds. Only 48% (n = 54) of hospitals report coverage of any newborn beds by pediatric hos-
pitalists; 21% (n = 23) report that hospitalists cover <5 newborn beds, 18% (n = 20) report coverage of 5 to 20 newborn beds, whereas just 10% (n = 11) cover >20 newborn beds.

In 80% (n = 90) of hospitals, hospitalists serve as the attending of record for all patients for whom they provide care. When asked about backup for pediatric hospitalists, respondents reported the subspecialist attending (46%; n = 52) and critical care attending (45%; n = 50) more often than either the emergency department (25%; n = 28) or the patient’s private physician (15%; n = 17). Non-COTH hospitals were more likely than COTH to have the emergency department provide back up (57% vs 6%; P = .0001).

A major role served by pediatric hospitalists is to communicate with primary care or referring physicians (PCPs). Nearly all hospitals (87%; n = 97) indicated that there is a defined or expected pattern for this communication. Common communication patterns stated by hospitals include: contact of PCP on admission (61%; n = 68); updating the PCP of progress during the patient’s course of stay (54%; n = 60); providing a discharge summary (71%; n = 80); and contacting the PCP on discharge (43%; n = 48).

Teaching Roles
The majority (73%; n = 82) of surveyed hospitals reported having ≥1 residency training program; 60% (n = 67) had pediatric residency programs (Table 4), whereas 45% (n = 50) and 41% (n = 46) reported training programs in internal medicine and family practice, respectively.

In programs with residents (n = 82), hospitalists serve as teaching attendings for pediatric patients in almost all cases (89%). Only 18% (n = 15) of these respondents reported that their hospital had separate teaching and nonteaching hospitalist services for pediatric patients. Most hospitalist program directors (89%; n = 73) believe that hospitalists are more effective teachers than rotating subspecialist attendings.

Employment Characteristics and Arrangements
Over half of hospitals (52%; n = 58) reported no turn-over in their hospitalists over the past 2 years. Pediatric hospitalist programs in existence for <2 years (n = 28) were excluded from the remainder of analyses on employment duration. When asked about the average duration of employment for a hospitalist at their institution, 46% of respondents stated that most were employed for <3 years; 29% indicated the average length of employment was 3 to 5 years, and 25% stated their hospitalists averaged >5 years on the job.

Most programs (59%; n = 66) reported that their hospitalists were employed by the hospital itself. Other employers of hospitalists included a university or medical school (25%; n = 28), general or specialty physician group (21%; n = 23), hospitalist-only group (8%; n = 9), or other (8%; n = 9).

With regard to compensation, most (72%; n = 81) reported that their hospitalists are 100% salaried. There were no significant differences across hospital types. Only 22% (n = 25) of respondents indicated that hospitalist compensation is at least partially dependent on other measures or incentives, such as patient satisfaction, quality assurance, controlling costs, or reducing length of stay.

For the 26% (n = 29) of hospitals that indicated there was a productivity component to their hospitalists’ compensation, common productivity requirements or standards included billings (n = 15) and relative value unit (n = 11).

The majority of hospitals surveyed reported that hospitalists do not generate enough income from professional billing to pay their salaries. In 70% (n = 78) of cases, respondents indicated that supplemental funding was required. Sources of funding include the hospital (n

---

**TABLE 3** Number of Pediatric Hospitalists per Hospital (N = 112)

<table>
<thead>
<tr>
<th>No. of Hospitalists</th>
<th>Original Total, % (n)</th>
<th>COTH, %</th>
<th>Non-COTH, %</th>
<th>NACHRI, %</th>
<th>Non-NACHRI, %</th>
<th>Freestanding, %</th>
<th>Nonfreestanding, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–5</td>
<td>59 (66)</td>
<td>53</td>
<td>68</td>
<td>62</td>
<td>65</td>
<td>21</td>
<td>67</td>
</tr>
<tr>
<td>6–10</td>
<td>30 (34)</td>
<td>36</td>
<td>32</td>
<td>25</td>
<td>33</td>
<td>49</td>
<td>32</td>
</tr>
<tr>
<td>&gt;10</td>
<td>11 (12)</td>
<td>12</td>
<td>0.49</td>
<td>13</td>
<td>1</td>
<td>30</td>
<td>2</td>
</tr>
</tbody>
</table>

P = .0019 .0013 <.0001

**TABLE 4** Hospitals With Residency Training Programs in Pediatrics (N = 82)

<table>
<thead>
<tr>
<th>Original Total, % (n)</th>
<th>COTH, %</th>
<th>Non-COTH, %</th>
<th>NACHRI, %</th>
<th>Non-NACHRI, %</th>
<th>Freestanding, %</th>
<th>Nonfreestanding, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>83 (67)</td>
<td>82</td>
<td>48</td>
<td>92</td>
<td>54</td>
<td>96</td>
</tr>
<tr>
<td>No</td>
<td>17 (14)</td>
<td>18</td>
<td>52</td>
<td>8</td>
<td>46</td>
<td>4</td>
</tr>
</tbody>
</table>

P = 1.248 .012 .0212

Frequency missing = 1.
Medical school, or department of pediatrics (n = 49). Significant differences in funding sources were demonstrated for both teaching and children’s hospitals (Table 5).

Training
The majority of the pediatric hospitalists are pediatric-trained physicians; 73% (n = 82) report that 100% of their hospitalists providing care to children are pediatricians. Some respondents reported medicine-pediatric trained physicians (15%; n = 17), internists (2%; n = 2), critical care (7%; n = 8), and emergency physicians (2%; n = 2) serving as pediatric hospitalists, but none reported family physicians in this role.

Requirements for Hospitalists Who Care for Pediatric Patients
Most programs (82%; n = 92) state they require board certification in pediatrics, and 46% (n = 51) require pediatric advanced life support certification. In addition, most hospitals (76%; n = 85) do not require any period of mandatory supervision before a hospitalist can work independently. Of the 21 hospitals that do require a supervision period, 13 require ≤1 month. In most cases (n = 13), a senior hospitalist provides the supervision. Nearly all hospitals (95%; n = 106) do not require hospitalists to complete continuing medical education or training specifically related to hospital medicine.

Outcomes Associated With Hospitalists
Fewer than half (39%; n = 44) of respondents reported that their hospital measures pediatric clinical outcomes associated with hospitalist care. Of these 44 hospitals, the most common outcomes measured include length of stay (n = 39), readmission rates (n = 35), condition-specific process measures (n = 29), quality of care measures (n = 25), mortality (n = 24), timing of discharge (n = 21), and quality of transfers (n = 10). Clinical outcomes are measured significantly more often in non-COTH (compared with COTH) and non-NACHRI (compared with NACHRI) hospitals (Table 6).

Self-Perceived Impact of Hospitalists
The vast majority (88%; n = 98) of respondents believe that use of a hospitalist service reduces hospital costs, whereas even more (92%; n = 103) believe that use of a hospitalist service increases patient satisfaction; 99% (n = 111) believe that use of a hospitalist service increases quality of care for inpatients. Finally, 94% (n = 105) do not agree with the statement that use of a hospitalist service adversely affects the primary care physician–patient relationship.

DISCUSSION
Our study provides the most comprehensive data to date on the emerging group of pediatric hospitalists. There was significant effort on both the part of hospitalist program directors and their professional associations to provide organizational structures for this group of clinicians. However, our finding that 46% of hospitalist program directors reported that most of their hospitalists had an average duration of employment of <3 years indicates there is a significant amount of turnover in this

### TABLE 5 Sources of Supplemental Funding (N = 78)

<table>
<thead>
<tr>
<th>Source of Funding</th>
<th>Original Total, % (n)</th>
<th>COTH, %</th>
<th>Non-COTH, %</th>
<th>NACHRI, %</th>
<th>Non-NACHRI, %</th>
<th>Freestanding, %</th>
<th>Nonfreestanding, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospital</td>
<td>63 (49)</td>
<td>50</td>
<td>80</td>
<td>55</td>
<td>78</td>
<td>66</td>
<td>75</td>
</tr>
<tr>
<td>Medical School Department of Pediatrics</td>
<td>17 (13)</td>
<td>27</td>
<td>1</td>
<td>28</td>
<td>2</td>
<td>18</td>
<td>5</td>
</tr>
<tr>
<td>Combination</td>
<td>12 (9)</td>
<td>18</td>
<td>1</td>
<td>16</td>
<td>2</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>Other</td>
<td>8 (6)</td>
<td>5</td>
<td>18</td>
<td>2</td>
<td>18</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>P</td>
<td>&lt;.0001</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

— indicates no value. Frequency missing = 1.

### TABLE 6 Pediatric Clinical Outcomes Measured for Hospitalists (N = 46)

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Original, % (n)</th>
<th>COTH, %</th>
<th>Non-COTH, %</th>
<th>NACHRI, %</th>
<th>Non-NACHRI, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>52 (24)</td>
<td>32</td>
<td>90</td>
<td>47</td>
<td>87</td>
</tr>
<tr>
<td>P</td>
<td></td>
<td>.0006</td>
<td>.0028</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Readmit rates</td>
<td>76 (35)</td>
<td>53</td>
<td>99</td>
<td>70</td>
<td>96</td>
</tr>
<tr>
<td>P</td>
<td></td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of care</td>
<td>56 (25)</td>
<td>34</td>
<td>97</td>
<td>44</td>
<td>94</td>
</tr>
<tr>
<td>P</td>
<td></td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition-specific process measures</td>
<td>63 (29)</td>
<td>44</td>
<td>90</td>
<td>61</td>
<td>87</td>
</tr>
<tr>
<td>P</td>
<td></td>
<td>.0059</td>
<td>.0336</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length of stay</td>
<td>85 (39)</td>
<td>71</td>
<td>99</td>
<td>84</td>
<td>97</td>
</tr>
<tr>
<td>P</td>
<td></td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

— indicates no value.
professional activity. The implications of this finding regarding the stability, and further definition of this specific group of clinicians are unclear but suggest this is a heterogeneous group of clinicians. Thus, hospitalist leaders should take note that there may be several subsets of individuals who are engaged in this line of work, including those who intend a long-term career in hospital care and those who do not.

Although there have been several studies suggesting that hospitalists may have a significant impact on clinical and financial outcomes, few hospitals actually have systems in place to measure such variables. Surprisingly, we found that only 39% measure any clinical outcomes associated with hospitalist care and only a subset of those measure length of stay or readmission rates. Until greater attention is focused on such issues, any actual benefit of hospitalists in these areas will remain theoretical and unproven.

The finding that 97 hospitals in our sampling frame had hospitalists caring for adults but not for children implies hospitals are using this model of care more frequently for their adult patients. Whether this trend will continue is unknown.

Likely because of their greater numbers, there is more research among internal medicine hospitalist programs. However, although there are several studies of hospitalist program outcomes in the internal medicine literature, almost all are single-site assessments and thus are likely subject to the nuances of local issues. Most such studies demonstrate a benefit from hospitalists with regard to lower costs and length of stay, but some fail to show any significant differences with previous inpatient systems. There are fewer published studies regarding the clinical impact of pediatric hospitalists, but they also are single site assessments and report mixed results.

Some have questioned whether a specific set of skills are required to practice pediatric hospital medicine. It is noteworthy that the hospitals providing employment to this group currently have relatively minimal standards for certification and training of hospitalists. For example, only 55% of hospitals require pediatric hospitalists to have completed residency training in pediatrics and only 46% require them to be certified in pediatric advanced life support. In addition, only 5% require any continuing medical education or training related specifically to hospital medicine. Thus, it is not surprising that a recent study found that fewer than 30% of pediatric academic department chairs believed that hospitalists require training not currently provided in residency.

It is unclear as to whether hospitalists are being hired by hospitals as part of a strategic decision to address quality, educational, and other issues, or whether they are employed to predominantly fill a need in clinical care. Our finding that >70% of hospitalist programs require supplemental funding to meet salary costs indicates institutions are actively investing in these programs. This could be because the cost savings attributed to hospitalist services would not translate into a greater volume of billing, but rather savings to their employing institutions in capitated or other incentive-based contracting arrangements. The payor mix of hospitalists in some institutions may also play a role in their ability to generate their salary costs. Some teaching hospitals may also be realizing a greater clinical service need after implementation of the 80-hour workweek for residents.

Our data clearly indicate that the leadership of hospitalist programs believes strongly that hospitalists bring value to both the clinical and educational arenas. Unfortunately, there is a paucity of generalizable data in the pediatric education literature regarding hospitalists, with all studies being single-site reports. In a recent review of this literature, Carlson et al note that “most studies on the impact of the hospitalist on pediatric medical education have been positive”; however, 1 study did report decreased ratings in resident autonomy and supervisory skills after the introduction of a hospitalist system. Clearly, a more comprehensive assessment of the educational impact on hospitalists nationwide is needed.

In 1999, the American Academy of Pediatrics recognized pediatric hospitalists by creating a Provisional Section on Hospital Care. In 2005, the American Academy of Pediatrics also issued a policy statement providing 6 guiding principles for the development of pediatric hospitalist programs. Our results suggest that at least some of these principles are not being implemented uniformly by current hospitalist programs including: (1) hospitalists should be board certified in pediatrics or have equivalent qualifications; (2) pediatric hospitalist programs should include appropriate outpatient follow-up; (3) pediatric hospitalist programs should provide for timely/complete communication between the hospitalist and the outpatient physician; and (4) pediatric hospitalist programs should include data collection and outcome assessment to monitor performance.

CONCLUSIONS

Although these findings are an important step in better understanding the current situation in pediatric hospital medicine, they provide only 1 perspective, that of hospitalist program directors. Additional perspectives specific to different stakeholders are essential for a more complete picture of this domain. For example, understanding both the motivations and the actual career plans of hospitalists will require studies of a representative sample of hospitalists themselves, not just their leadership. In the same vein, determination of the actual benefits perceived by a hospital in subsidizing a hospitalist program would be best learned from a sample of hospital chief executives or chief operating officers. Finally, an important perspective on the educational impact and utility of pediatric hospitalists will come from those entrusted with the integrity of our training pro-
grams, the residency program directors at each institution.

Hospital medicine is a rapidly growing enterprise. The better we understand its participants, as well as those affected by its practice, the better we will be able to plan for a future, meeting as many needs as possible while ensuring the best possible care for children.

ACKNOWLEDGMENT
This work was supported by the American Board of Pediatrics Foundation.

REFERENCES
17. Rogers JC. Pediatric hospitalist programs offer chance to improve quality and cost. Health Care Strateg Manage. 2003;21(3):12–15
Growth and Neurodevelopmental Outcomes After Early Low-Dose Hydrocortisone Treatment in Extremely Low Birth Weight Infants

Kristi L. Watterberg, MD, Michelle L. Shaffer, PhD, Mary J. Mishefske, MD, Corinne L. Leach, MD, PhD, Mark C. Mammel, MD, Robert J. Couser, MD, Soraya Abbasi, MD, Cynthia H. Cole, MD, Susan W. Aucott, MD, Elizabeth H. Thilo, MD, Henry J. Rozycki, MD, Conra Backstrom Lacy, RN

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

ABSTRACT

BACKGROUND. Low cortisol concentrations in premature infants have been correlated with increased severity of illness, hypotension, mortality, and development of bronchopulmonary dysplasia. A total of 360 mechanically ventilated infants with a birth weight of 500 to 999 g were enrolled in a randomized, multicenter trial of prophylaxis of early adrenal insufficiency to prevent bronchopulmonary dysplasia. Mortality and bronchopulmonary dysplasia were decreased in the hydrocortisone-treated patients exposed to chorioamnionitis. We now report outcomes at 18 to 22 months’ corrected age.

PATIENTS AND METHODS. Surviving infants were evaluated with standardized neurologic examination and Bayley Scales of Infant Development-II. Neurodevelopmental impairment was defined as a Mental Development Index or Psychomotor Developmental Index of <70, cerebral palsy, blindness or deafness.

RESULTS. A total of 252 (87%) of 291 survivors were evaluated. Cerebral palsy was diagnosed in 13% of hydrocortisone-treated versus 14% of placebo-treated infants. Fewer hydrocortisone-treated infants had a Mental Development Index <70, and more of the hydrocortisone-treated infants showed evidence of awareness of object permanence. Incidence of neurodevelopmental impairment was not different (39% [hydrocortisone] vs 44% [placebo]). There were no differences in physical growth measures. Chorioamnionitis-exposed infants treated with hydrocortisone were shorter and weighed less than controls but had no evidence of neurodevelopmental impairment. Among infants not exposed to chorioamnionitis, hydrocortisone-treated patients were less likely to have a Mental Development Index of <70 or to be receiving glucocorticoids at follow-up.

CONCLUSIONS. Early, low-dose hydrocortisone treatment was not associated with increased cerebral palsy. Treated infants had indicators of improved developmental outcome. Together with the short-term benefit previously reported, these data support additional studies of hydrocortisone treatment of adrenal insufficiency in extremely premature infants.
Extremely premature infants are at high risk for numerous adverse outcomes, including death, bronchopulmonary dysplasia (BPD), and neurodevelopmental impairment (NDI).1 2 Early dexamethasone therapy has been shown to prevent or decrease the severity of BPD in preterm infants; however, both short-term and long-term adverse effects have limited its use.3 Specifically, dexamethasone therapy has been associated with an increase in cerebral palsy (CP) and other NDIs at follow-up.2 3–4 Low cortisol values and decreased response to corticotropin stimulation in these infants have been associated with increased severity of illness, increased mortality, and subsequent development of BPD.5–9 We developed the hypothesis that inadequate adrenal function in the face of critical illness in such infants contributes to adverse outcomes including BPD, and that prophylaxis of early adrenal insufficiency in these infants would improve survival without BPD. A pilot study supported that hypothesis10 and led to the development of a multicenter, randomized trial of prophylaxis of early adrenal insufficiency to prevent BPD in intubated, extremely low birth weight (ELBW) infants.

This recently completed multicenter, randomized trial showed no overall improvement in survival without BPD for hydrocortisone-treated infants; however, treated infants exposed to chorioamnionitis had significantly increased survival and survival without BPD.11 Enrollment in the study was stopped at 360 patients, approximately half the planned enrollment, because of an increase in the incidence of spontaneous gastrointestinal perforation in the hydrocortisone-treated infants, likely because of an interaction with early indomethacin therapy.11 With the exception of the increase in gastrointestinal perforation, the hydrocortisone-treated infants did not experience any of the short-term adverse effects previously reported with early dexamethasone therapy. We are now reporting the outcomes of these infants at 18 to 22 months’ adjusted age.

METHODS

Population and Study Protocol

Infants eligible for this follow-up study were surviving infants who had been enrolled in the multicenter study of low-dose hydrocortisone therapy for prophylaxis of early adrenal insufficiency.11 Details of these methods have been previously reported. Briefly, singletons and twins between 500 and 999 g birth weight were eligible if they were mechanically ventilated via an endotracheal tube at study entry (12–48 hours of life). The study protocol was approved by institutional review boards at all participating institutions, and parental consent was obtained before enrollment. Randomization was stratified by study center and birth weight (500–749 and 750–999 g). Twins were randomized together to the same study arm. Infants received normal saline placebo or hydrocortisone sodium succinate (Solu-Cortef Plain, Amersham Pharmacia & Upjohn, Kalamazoo, MI), 1 mg/kg per day (~8–10 mg/m² per day), divided twice daily for 12 days, followed by 0.5 mg/kg per day for 3 days. Because the pilot study demonstrated particular benefit for infants exposed to chorioamnionitis, this group was of specific a priori interest; therefore, all placental histology was reviewed and graded by 2 central readers (Nancy Joste, MD, and Marcia Wills, MD).

Follow-up Study Procedures

At the follow-up visit, demographic and medical histories were obtained. Weight, height, and head circumference were recorded. Growth outcomes were adjusted for corrected age using z scores on the basis of the Centers for Disease Control and Prevention 2000 growth charts.12 Before the follow-up phase of this study, the neurologic examiners met and agreed on a standardized neurologic examination, with specific definitions for each component. CP 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. Functional gross motor level was assessed by using a standardized 5-level classification system.13 Development was assessed with the Bayley Scales of Infant Development II.14 A Mental Development Index (MDI) or Psychomotor Development Index (PDI) >2 SD below the mean (ie, <70) was defined as abnormal. For children who scored <50, a score of 49 was assigned. Behavior was assessed with the Bayley Behavior Rating Scale (BRS). Bayley mental scale items 84, 96, and 102 were predetermined as measures of object permanence to assess prefrontal cortex development. The children were asked to find a toy hidden under 1 of 2 cups, with double visual displacement used to increase the difficulty of the item. All Bayley examiners had been previously certified, and each examiner submitted a scored tape to 1 central examiner (Jean Lowe, PhD) who reviewed the tapes for consistency. Study assignment remained masked throughout the follow-up period, and no examiner was aware of the treatment assignment of any infant.

NDI was defined as at least 1 of the following: CP, MDI <70, PDI <70, functional deafness, or functional blindness. Functional deafness was defined as the inability to successfully complete the MDI, PDI, or BRS because of an auditory sensory impairment. Functional blindness was defined as the inability to successfully complete the MDI, PDI, or BRS because of a visual sensory impairment.

Statistical Analysis

Because of the previously reported association of dexamethasone with CP in ELBW infants, the study was powered to detect an increase of 10 percentage points
(1-sided hypothesis test) in the outcome of CP for hydrocortisone-treated infants compared with placebo-treated infants. A sample size of 712 births (including eligible second twins, the anticipated sample size was 790 infants) was required to achieve a power of .80 to detect the 10% increase with \( \alpha = .05 \), assuming a survival rate of 85%, follow-up rate of \( \geq 80\% \), and an incidence of CP \( \geq 20\% \) in the placebo group. Because study enrollment was stopped at 360 infants, statistically insignificant results must be viewed with caution because of the increased probability of a type II error.

All infants evaluated for long-term outcomes were included in an intent-to-treat analysis. Baseline characteristics, 36-week outcomes, and population characteristics at follow-up for the treatment groups were compared by using 2-sample \( t \) tests for continuous outcomes and \( \chi^2 \) or Fisher’s exact tests for categorical outcomes. Neurodevelopmental outcomes were analyzed by using analysis of covariance for continuous outcomes and logistic regression for binary outcomes. These analyses included adjustment for the stratification variables birth weight (continuous form) and center and the following baseline characteristics: gestational age, prenatal steroid use, outborn, gender, black race, and method of delivery. These maximally adjusted analyses were compared with analyses that included adjustment for only birth weight and gestational age where overfitting was a concern. Similar analyses were conducted for 36-week outcomes and population characteristics at follow-up by using analysis of covariance and logistic regression to examine the impact of stratification variables and risk factors. Unless otherwise noted, all hypotheses tests were 2-sided and used a significance level of .05.

**RESULTS**

A total of 360 infants were enrolled in this study at 9 study centers between November 2001 and April 2003; 294 of these remained in the study and survived to discharge (Fig 1). Of these, 3 were known to have died before follow-up, leaving 291 available for follow-up. Of these, 252 (87%) were evaluated for long-term outcomes. All baseline characteristics were similar between infants seen in follow-up and those lost to follow-up with the exception of ethnicity. A higher percentage of the children lost to follow-up were Hispanic (10 of 39 lost to follow-up were Hispanic versus 25 of 252 children seen in follow-up).

Table 1 shows the patient characteristics at study entry and outcomes at 36 weeks’ postmenstrual age for those infants seen at follow-up. Hydrocortisone-treated and placebo infants were similar, with the exception of gastrointestinal perforation, as previously reported.\(^{11}\) Tables 2–4 show the characteristics of these infants at the follow-up examination, presented first for the entire study population, then for patients known to have or not have chorioamnionitis on placental histologic examination. Maternal education and household income level were similar between all groups. For the overall study groups, there were no significant differences between the hydrocortisone-treated infants and those who received placebo. In the subset of infants exposed to chorioamnionitis, hydrocortisone-treated infants had significantly lower weight and height than the placebo infants. In the subset of infants not exposed to chorioamnionitis, significantly fewer hydrocortisone-treated infants were being treated with inhaled or systemic corticosteroids.

Tables 5–7 present the neurodevelopmental outcomes. The primary outcome variable chosen for the study, CP, was not different between groups. Comparing the overall study groups, significantly fewer hydrocortisone-treated patients had a Bayley MDI of <70 (\( P = .017 \)), and more of the hydrocortisone-treated children showed evidence of awareness of object permanence on the Bayley examination (\( P = .035 \)). Other measures were not significantly different between the groups, although the direction of effect favored the hydrocortisone-treated infants. For chorioamnionitis-exposed infants, there were no differences between groups, with a trend toward more independent feeding in the hydrocortisone-treated infants (\( P = .056 \)). For infants not exposed to chorioamnionitis, fewer hydrocortisone-treated infants had a Bayley MDI of <70 (\( P = .025 \)), and the direction of all differences favored the hydrocortisone-treated infants. There were no differences in the BRS results between groups.

Fourteen of 17 surviving infants who experienced spontaneous gastrointestinal perforation were evaluated. Because of the small size of this group, statistical comparisons would not be meaningful; however, their
outcomes seemed to be similar to the remainder of the patients (eg, NDI: 43% vs 41%; weight: 10.8 vs 10.8 kg).

### DISCUSSION

In this multicenter, placebo-controlled, randomized trial of early, low-dose hydrocortisone treatment for prophylaxis of early adrenal insufficiency, we found that infants treated with hydrocortisone did not have an increase in CP or other evidence of NDI at follow-up compared with the placebo group, in contrast to previously reported effects of higher-dose dexamethasone. Although patient enrollment was stopped early, decreasing the power to detect a difference, we found no trend toward such outcomes. To the contrary, hydrocortisone-treated infants showed evidence of improved neurologic outcome, with a significantly lower incidence of Bayley MDI of <70 and better awareness of object permanence. Interestingly, although hydrocortisone treatment improved the short-term outcomes of survival and survival without BPD only for infants exposed to chorioamnionitis, the infants not exposed to chorioamnionitis seemed to derive

---

**TABLE 1** Baseline Population Characteristics and Short-term Outcomes of Patients Seen in Follow-up

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Hydrocortisone-Treated Group</th>
<th>Placebo Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N = 126)</td>
<td>(N = 126)</td>
</tr>
<tr>
<td>Birth weight, mean ± SD, g</td>
<td>738 ± 122</td>
<td>738 ± 124</td>
</tr>
<tr>
<td>Gestation, mean ± SD, wk</td>
<td>25.3 ± 1.5</td>
<td>25.4 ± 1.6</td>
</tr>
<tr>
<td>Head circumference (birth), mean ± SD, cm</td>
<td>22.8 ± 1.6</td>
<td>23.0 ± 1.4</td>
</tr>
<tr>
<td>Male gender, n (%)</td>
<td>68 (54)</td>
<td>64 (51)</td>
</tr>
<tr>
<td>Outborn, n (%)</td>
<td>17 (13)</td>
<td>9 (7)</td>
</tr>
<tr>
<td>Racial/ethnic group, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>68 (54)</td>
<td>50 (40)</td>
</tr>
<tr>
<td>Black</td>
<td>41 (33)</td>
<td>51 (40)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>12 (10)</td>
<td>13 (10)</td>
</tr>
<tr>
<td>Other</td>
<td>5 (4)</td>
<td>12 (10)</td>
</tr>
<tr>
<td>Chorioamnionitis, n/N (%)</td>
<td>57/103 (55)</td>
<td>55/107 (51)</td>
</tr>
</tbody>
</table>

36-wk outcomes

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Hydrocortisone-Treated Group</th>
<th>Placebo Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight, mean ± SD, g</td>
<td>2008 ± 299</td>
<td>2047 ± 346</td>
</tr>
<tr>
<td>Head circumference, mean ± SD, cm</td>
<td>31.2 ± 1.5</td>
<td>31.0 ± 1.5</td>
</tr>
<tr>
<td>BPD (clinical), n (%)</td>
<td>74 (59)</td>
<td>76 (60)</td>
</tr>
<tr>
<td>BPD (physiologic), n/N (%)</td>
<td>57/118 (48)</td>
<td>54/111 (49)</td>
</tr>
<tr>
<td>Gastrointestinal perforation, n (%)</td>
<td>13 (10)</td>
<td>1 (1)*</td>
</tr>
<tr>
<td>Grade III/IV intracranial hemorrhage, n/N (%)</td>
<td>17/123 (14)</td>
<td>13/125 (10)</td>
</tr>
<tr>
<td>Periventricular leukomalacia, n/N (%)</td>
<td>7/99 (7)</td>
<td>8/105 (8)</td>
</tr>
<tr>
<td>Surgery for retinopathy of prematurity, n (%)</td>
<td>17 (13)</td>
<td>17 (13)</td>
</tr>
</tbody>
</table>

BPD (clinical) indicates receiving supplemental oxygen at 36 weeks' postmenstrual age; BPD (physiologic), supplemental oxygen was required to maintain oxygen saturations >89% during a 4-hour observation period.

* Significantly different between groups: P < .01.

**TABLE 2** Population Characteristics at Follow-up

<table>
<thead>
<tr>
<th>Characteristic at Follow-up</th>
<th>Hydrocortisone-Treated Group</th>
<th>Placebo Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N = 126)</td>
<td>(N = 126)</td>
</tr>
<tr>
<td>Adjusted age, mean ± SD, mo</td>
<td>20.0 ± 2.1</td>
<td>20.0 ± 2.1</td>
</tr>
<tr>
<td>Maternal education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school graduate or less</td>
<td>49</td>
<td>53</td>
</tr>
<tr>
<td>Some college/trade school</td>
<td>41</td>
<td>35</td>
</tr>
<tr>
<td>At least a college degree</td>
<td>30</td>
<td>34</td>
</tr>
<tr>
<td>Unknown</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Household income, median</td>
<td>$30,000–$40,000</td>
<td>$30,000–$40,000</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>10.6 ± 2.1</td>
<td>10.9 ± 1.5</td>
</tr>
<tr>
<td>z score</td>
<td>−1.08 ± 1.4</td>
<td>−0.80 ± 1.3</td>
</tr>
<tr>
<td>Height, cm</td>
<td>80.7 ± 3.5</td>
<td>81.5 ± 4.0</td>
</tr>
<tr>
<td>z score</td>
<td>−0.67 ± 1.0</td>
<td>−0.42 ± 1.1</td>
</tr>
<tr>
<td>Head circumference, cm</td>
<td>47.1 ± 1.9</td>
<td>47.3 ± 1.8</td>
</tr>
<tr>
<td>z score</td>
<td>−0.28 ± 1.3</td>
<td>−0.09 ± 1.3</td>
</tr>
<tr>
<td>Current therapies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxygen, n (%)</td>
<td>4 (3)</td>
<td>7 (6)</td>
</tr>
<tr>
<td>Bronchodilators, n (%)</td>
<td>21 (17)</td>
<td>29 (23)</td>
</tr>
<tr>
<td>Steroids, systemic/inhaled (total %)</td>
<td>2/21 (10)</td>
<td>3/29 (25)</td>
</tr>
<tr>
<td>Rehospitalized after discharge, n (%)</td>
<td>65 (52)</td>
<td>67 (53)</td>
</tr>
</tbody>
</table>
significant benefit in neurodevelopmental and medical outcomes at 18 to 22 months’ corrected age. Interpretation of these subgroup findings, although planned a priori, should be made with caution, because the numbers are small.

Few studies have investigated the use of hydrocortisone for treatment or prevention of BPD in premature infants, and even fewer have described long-term outcomes after neonatal hydrocortisone treatment. A small, retrospective cohort study comparing hydrocortisone with dexamethasone suggested that hydrocortisone may be as effective as dexamethasone in reducing BPD, with fewer immediate and long-term adverse effects. Those authors subsequently reported structural and functional brain development at 8 years of age in their patients, finding that hydrocortisone-treated infants had intelligence scores and MRI findings similar to a cohort of preterm infants not treated with hydrocortisone, although the hydrocortisone-treated infants were significantly smaller, more immature, and sicker than infants not treated with postnatal glucocorticoids. In contrast, infants treated with dexamethasone were reported to have adverse developmental outcomes, as well as impaired cerebral gray matter growth on MRI.

Both our findings and those of the other studies summarized above are consistent with previously described effects of cortisol and synthetic glucocorticoids in the brain. Using animal models, investigators have delineated an inverted U pattern for cortisol or corticosterone effects on the central nervous system, such that both

### TABLE 3  Population Characteristics at Follow-up: Patients With Chorioamnionitis

<table>
<thead>
<tr>
<th>Characteristic at Follow-up</th>
<th>Hydrocortisone-Treated Group</th>
<th>Placebo Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N = 57)</td>
<td>(N = 55)</td>
</tr>
<tr>
<td>Adjusted age, mean ± SD, mo</td>
<td>19.3 ± 1.7</td>
<td>20.0 ± 2.0</td>
</tr>
<tr>
<td>Maternal education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school graduate or less</td>
<td>20</td>
<td>26</td>
</tr>
<tr>
<td>Some college/trade school</td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td>At least a college degree</td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>Unknown</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Household income, median</td>
<td>$30,000–$40,000</td>
<td>$30,000–$40,000</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>10.4 ± 1.4</td>
<td>11.2 ± 1.3</td>
</tr>
<tr>
<td>z score</td>
<td>−1.14 ± 1.3</td>
<td>−0.60 ± 1.1</td>
</tr>
<tr>
<td>Height, cm</td>
<td>80.1 ± 3.6</td>
<td>82.5 ± 3.9</td>
</tr>
<tr>
<td>z score</td>
<td>−0.71 ± 1.1</td>
<td>−0.20 ± 1.1</td>
</tr>
<tr>
<td>Head circumference, cm</td>
<td>46.8 ± 1.7</td>
<td>47.5 ± 1.7</td>
</tr>
<tr>
<td>z score</td>
<td>−0.45 ± 1.1</td>
<td>−0.04 ± 1.3</td>
</tr>
<tr>
<td>Current therapies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxygen, n (%)</td>
<td>3 (5)</td>
<td>5 (9)</td>
</tr>
<tr>
<td>Bronchodilators, n (%)</td>
<td>7 (12)</td>
<td>13 (24)</td>
</tr>
<tr>
<td>Steroids, systemic/inhaled (total %)</td>
<td>2/9 (19)</td>
<td>1/10 (20)</td>
</tr>
<tr>
<td>Rehospitalized after discharge, n (%)</td>
<td>30 (53)</td>
<td>28 (51)</td>
</tr>
</tbody>
</table>

*Significantly different between groups: P < .05.

### TABLE 4  Population Characteristics at Follow-up: Patients Without Chorioamnionitis

<table>
<thead>
<tr>
<th>Characteristic at Follow-up</th>
<th>Hydrocortisone-Treated Group</th>
<th>Placebo Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N = 46)</td>
<td>(N = 52)</td>
</tr>
<tr>
<td>Adjusted age, mean ± SD, mo</td>
<td>20.3 ± 2.3</td>
<td>20.1 ± 2.5</td>
</tr>
<tr>
<td>Maternal education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school graduate or less</td>
<td>18</td>
<td>13</td>
</tr>
<tr>
<td>Some college/trade school</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td>At least a college degree</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>Unknown</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Household income, median</td>
<td>$40,000–$50,000</td>
<td>$40,000–$50,000</td>
</tr>
<tr>
<td>Weight, mean ± SD, kg</td>
<td>10.4 ± 1.4</td>
<td>10.6 ± 1.6</td>
</tr>
<tr>
<td>z score</td>
<td>−1.25 ± 1.2</td>
<td>−1.04 ± 1.3</td>
</tr>
<tr>
<td>Height, cm</td>
<td>80.7 ± 3.3</td>
<td>80.9 ± 4.1</td>
</tr>
<tr>
<td>z score</td>
<td>−0.74 ± 1.0</td>
<td>−0.64 ± 1.1</td>
</tr>
<tr>
<td>Head circumference, cm</td>
<td>47.4 ± 1.8</td>
<td>47.4 ± 1.8</td>
</tr>
<tr>
<td>z score</td>
<td>−0.10 ± 1.3</td>
<td>−0.03 ± 1.3</td>
</tr>
<tr>
<td>Current therapies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxygen, n (%)</td>
<td>1 (2)</td>
<td>2 (4)</td>
</tr>
<tr>
<td>Bronchodilators, n (%)</td>
<td>8 (18)</td>
<td>15 (29)</td>
</tr>
<tr>
<td>Steroids, systemic/inhaled (total %)</td>
<td>0/4 (9)</td>
<td>2/17 (37)</td>
</tr>
</tbody>
</table>

*Significantly different between groups: P = .002.
very low and very high concentrations are associated with adverse central nervous system effects.\(^1\) Thus, although sustained excessive cortisol concentrations produce detrimental effects, particularly in the hippocampus, adrenalectomy also adversely affects structure and function.\(^2\) Cortisol occupies both mineralocorticoid receptors and glucocorticoid receptors in the brain, binding preferentially to mineralocorticoid receptors at normal physiologic concentrations.\(^1\) Dexamethasone, however, binds only to glucocorticoid receptors. For that reason, and also because of its limited transmission into the brain, it was postulated that dexamethasone exerts

### TABLE 5  Neurodevelopmental Outcomes

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Hydrocortisone-Treated Group ((N = 126))</th>
<th>Placebo Group ((N = 126))</th>
<th>Odds Ratio (95% Confidence Interval) or (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP, (n) (%)</td>
<td>16 (13)</td>
<td>18 (14)</td>
<td>0.71 (0.33–1.57)</td>
</tr>
<tr>
<td>Bayley MDI, mean ± SD</td>
<td>80 ± 19</td>
<td>77 ± 19</td>
<td>0.8</td>
</tr>
<tr>
<td>Bayley PDI, mean ± SD</td>
<td>83 ± 19</td>
<td>84 ± 20</td>
<td>0.82</td>
</tr>
<tr>
<td>Bayley BRS, mean ± SD</td>
<td>36 ± 27</td>
<td>36 ± 27</td>
<td>0.76</td>
</tr>
<tr>
<td>Awareness of object permanence, %</td>
<td>89</td>
<td>79</td>
<td>2.19 (1.06–4.52)(^*)</td>
</tr>
<tr>
<td>Bayley MDI &lt; 70, %</td>
<td>27</td>
<td>37</td>
<td>0.47 (0.25–0.87)(^*)</td>
</tr>
<tr>
<td>Bayley PDI &lt; 70, %</td>
<td>26</td>
<td>23</td>
<td>1.03 (0.55–1.91)</td>
</tr>
<tr>
<td>NDI</td>
<td>39</td>
<td>44</td>
<td>0.66 (0.37–1.14)</td>
</tr>
<tr>
<td>NDI or death, (n/N) (%)</td>
<td>81/156 (52)</td>
<td>88/158 (56)</td>
<td>0.68 (0.41–1.10)</td>
</tr>
<tr>
<td>Normal functional gross motor, (n)</td>
<td>97 (77)</td>
<td>93 (74)</td>
<td>1.35 (0.74–2.51)</td>
</tr>
<tr>
<td>Normal gait, (n) (%)</td>
<td>90 (71)</td>
<td>84 (67)</td>
<td>1.36 (0.75–2.44)</td>
</tr>
<tr>
<td>Independent feeding, (n) (%)</td>
<td>90 (73)</td>
<td>82 (65)</td>
<td>1.72 (0.96–3.08)</td>
</tr>
</tbody>
</table>

Odds ratios and \(P\)-value analyses were adjusted for stratification variables as described in the text.

\(^*\) Significant difference between groups, \(P = .03\).

### TABLE 6  Neurodevelopmental Outcomes: Patients With Chorioamnionitis

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Hydrocortisone-Treated Group ((N = 57))</th>
<th>Placebo Group ((N = 55))</th>
<th>Odds Ratio (95% Confidence Interval) or (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP, (n) (%)</td>
<td>7 (12)</td>
<td>7 (13)</td>
<td>0.94 (0.24–3.71)</td>
</tr>
<tr>
<td>Bayley MDI, mean ± SD</td>
<td>81 ± 19</td>
<td>78 ± 18</td>
<td>0.97</td>
</tr>
<tr>
<td>Bayley PDI, mean ± SD</td>
<td>82 ± 19</td>
<td>86 ± 20</td>
<td>0.35</td>
</tr>
<tr>
<td>Bayley BRS, mean ± SD</td>
<td>34 ± 25</td>
<td>35 ± 25</td>
<td>0.68</td>
</tr>
<tr>
<td>Awareness of object permanence, %</td>
<td>90</td>
<td>77</td>
<td>2.77 (0.87–8.84)</td>
</tr>
<tr>
<td>Bayley MDI &lt; 70, %</td>
<td>30</td>
<td>31</td>
<td>1.09 (0.40–2.96)</td>
</tr>
<tr>
<td>Bayley PDI &lt; 70, %</td>
<td>30</td>
<td>22</td>
<td>1.73 (0.64–4.65)</td>
</tr>
<tr>
<td>NDI, %</td>
<td>40</td>
<td>40</td>
<td>1.22 (0.48–3.11)</td>
</tr>
<tr>
<td>NDI or death, (n/N) (%)</td>
<td>33/67 (49)</td>
<td>40/73 (55)</td>
<td>0.64 (0.29–1.42)</td>
</tr>
<tr>
<td>Normal functional gross motor, %</td>
<td>81</td>
<td>75</td>
<td>1.56 (0.53–4.61)</td>
</tr>
<tr>
<td>Normal gait, %</td>
<td>68</td>
<td>71</td>
<td>0.68 (0.26–1.80)</td>
</tr>
<tr>
<td>Independent feeding, %</td>
<td>77</td>
<td>60</td>
<td>2.57 (0.98–6.74)</td>
</tr>
</tbody>
</table>

Odds ratios and \(P\)-value analyses were adjusted for stratification variables as described in the text.

### TABLE 7  Neurodevelopmental Outcomes: Patients Without Chorioamnionitis

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Hydrocortisone-Treated ((N = 46))</th>
<th>Placebo ((N = 52))</th>
<th>Odds Ratio (95% Confidence Interval) or (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP, (n) (%)</td>
<td>6 (13)</td>
<td>8 (15)</td>
<td>0.75 (0.16–3.43)</td>
</tr>
<tr>
<td>Bayley MDI, mean ± SD</td>
<td>81 ± 20</td>
<td>77 ± 20</td>
<td>0.33</td>
</tr>
<tr>
<td>Bayley PDI, mean ± SD</td>
<td>84 ± 21</td>
<td>82 ± 20</td>
<td>0.49</td>
</tr>
<tr>
<td>Bayley BRS, mean ± SD</td>
<td>39 ± 32</td>
<td>36 ± 29</td>
<td>0.93</td>
</tr>
<tr>
<td>Awareness of object permanence, %</td>
<td>88</td>
<td>81</td>
<td>1.75 (0.55–5.62)</td>
</tr>
<tr>
<td>Bayley MDI &lt; 70, %</td>
<td>24</td>
<td>42</td>
<td>0.24 (0.07–0.84)(^*)</td>
</tr>
<tr>
<td>Bayley PDI &lt; 70, %</td>
<td>22</td>
<td>25</td>
<td>0.49 (0.12–1.99)</td>
</tr>
<tr>
<td>NDI, %</td>
<td>37</td>
<td>48</td>
<td>0.41 (0.14–1.24)</td>
</tr>
<tr>
<td>NDI or death, (n/N) (%)</td>
<td>28/55 (51)</td>
<td>32/59 (54)</td>
<td>0.52 (0.20–1.32)</td>
</tr>
<tr>
<td>Normal functional gross motor, %</td>
<td>76</td>
<td>71</td>
<td>1.76 (0.60–5.18)</td>
</tr>
<tr>
<td>Normal gait, %</td>
<td>76</td>
<td>60</td>
<td>1.90 (0.71–5.55)</td>
</tr>
<tr>
<td>Independent feeding, %</td>
<td>77</td>
<td>65</td>
<td>1.94 (0.67–5.56)</td>
</tr>
</tbody>
</table>

Odds ratios and \(P\)-value analyses adjusted for stratification variables as described in text.

\(^*\) Significant difference between groups, \(P = .025\).
its adverse effects on the hippocampus by causing a “chemical adrenalectomy.” Consistent with that hypothesis, administration of corticosterone to adrenalectomized adult rats was protective against the apoptotic effects of dexamethasone.

Although our findings that hydrocortisone did not increase the incidence of CP and instead conferred possible neurodevelopmental benefit must be confirmed in future trials, there are plausible mechanisms to explain these findings. First, the absence of neurodevelopmental harm likely resulted from the much lower dose of glucocorticoid administered; high doses of all glucocorticoids produce global growth impairment. In addition, the hydrocortisone preparation did not contain a sulfite preservative, which has been associated with adverse neurologic effects in animal models. The neurodevelopmental benefit may have resulted from improved cardiovascular function and better perfusion of the brain, and/or from direct interaction of hydrocortisone with the brain. One important such interaction is modulation of the immune response. In the rat model, corticosterone was shown to play a major role in controlling cerebral innate immunity, specifically suppressing microglial uptake of glutamate and production of tumor necrosis factor α. Administration of a glucocorticoid receptor inhibitor in that model leads to an amplified immune response to inflammatory stimulus and results in neurotoxicity, again suggesting that adrenal insufficiency is deleterious to the brain.

Similar to previous reports, the infants in this study who were exposed to chorioamnionitis had higher cortisol concentrations at study entry. Other studies have also shown that premature infants exposed to chorioamnionitis have both higher cortisol concentrations and increased inflammation early in life. Therefore, our finding of early benefit in regard to death and BPD in infants exposed to chorioamnionitis may have derived from the antiinflammatory effects of hydrocortisone in the lung. On the other hand, because infants not exposed to chorioamnionitis have lower cortisol concentrations in the first weeks of life, the long-term neurodevelopmental benefits seen in those infants may have derived from the effects of hydrocortisone therapy on adrenal insufficiency. Accumulating evidence supports the occurrence of relative adrenal insufficiency in very preterm infants. Lower cortisol values have been documented in infants with higher illness scores, those receiving vasopressor support, and those who die, as well as those who subsequently develop BPD. In addition, infants with vasopressor-resistant hypotension typically respond to administration of hydrocortisone, and hydrocortisone is being increasingly used for this purpose in extremely preterm infants.

In our study, we found that physical growth measures were not different in the overall study groups, but weight and length were lower in the group of hydrocortisone-treated patients exposed to chorioamnionitis compared with those in the placebo group. The reason for this is not clear, particularly because the hydrocortisone-treated children had no evidence of increased NDI, which would have been consistent with decreased physical growth, as reported for infants exposed to dexamethasone. These findings may represent a survivor effect, because significantly more hydrocortisone-treated infants survived than did the placebo infants in this group. Alternatively, among chorioamnionitis-exposed infants, the hydrocortisone-treated group had a mean birth weight 4% lower and head circumference 3% smaller than the placebo-treated infants at study entry, and those differences may have persisted at outcome. The finding could also be an artifact of a smaller than planned sample size.

The adverse event that caused early closure of the study, gastrointestinal perforation, seems likely to be because of an interaction with early indomethacin therapy in infants with high cortisol concentrations. To help balance risk and benefit, future studies may monitor cortisol concentrations to guide therapeutic decision making for individual patients.

**CONCLUSIONS**

We found that low-dose hydrocortisone therapy for prophylaxis of early adrenal insufficiency did not increase the incidence of CP and seemed to confer some neurodevelopmental benefit at 18 to 22 months’ corrected age in extremely preterm infants. This contrasts with dexamethasone, which has been shown to have adverse neurodevelopmental effects, prompting the American Academy of Pediatrics and the Canadian Paediatric Society to strongly caution against its use. Our findings, together with the growing evidence that low cortisol concentrations in this population correlate with adverse clinical manifestations, support additional randomized, controlled studies of low-dose hydrocortisone treatment in extremely preterm infants.

**ACKNOWLEDGMENTS**

This work was supported by a grant from the National Institute of Child Health and Human Development (R01–HD38540) and by grants from the General Clinical Research Centers Programs at the University of New Mexico (MO1 RRO0054), Tufts-New England Medical Center (5MO1 RRO0997), and the University of Colorado (MO1–RRO069).

The following were participating institutions, NICUs, and other investigators: University of New Mexico: Children’s Hospital of New Mexico NICU (Rebecca Montman, RN, Virginia Laadt, PhD, Jean Lowe, PhD, and Gerri Duran, MS) and University of New Mexico Department of Pathology (Nancy Joste, MD, and Marcia Wills, MD); Pennsylvania Hospital; Pennsylvania Hospital NICU (Jeremy Gerdes, MD, and Toni Mancini, RN); Tufts Univer-
REFERENCES


STUDY FINDS INCREASE IN BABY TEETH CAVITIES

“Tooth decay in young children’s baby teeth is on the rise, according to the largest government study of the nation’s dental health in more than 25 years. . . . ‘Overall, we can say that most Americans are noticing an improvement in their oral health,’ said the lead author, Dr Bruce Dye of the National Center for Health Statistics. The prevalence of cavities in baby teeth of children ages 2 to 5 increased to 28 percent in 1999–2004, from 24 percent in 1988–1994. Tooth decay in young children had been decreasing for 40 years. The new report contains the first statistically significant proof the trend has reversed, experts said. A reason is that parents are giving their children more processed snack foods than in the past, and more bottled water or other drinks instead of fluoridated tap water, Dr Dye said. Inadequate dental care may also play a role. Cavities in children can form quickly, and parents should begin bringing their children to the dentist at age 1, said Dr Joel Bergof of the University of Washington.”

Noted by JFL, MD
Botulinum Toxin for Spasticity in Children With Cerebral Palsy: A Comprehensive Evaluation

Kristie Bjornson, PhD, PT, a Ross Hays, MD, b Cathy Graubert, PT, a Robert Price, MS, b Francine Won, PT, a John F. McLaughlin, MD, c Morty Cohen, RPh a

aDepartment of Rehabilitation Medicine, Children’s Hospital and Regional Medical Center, Seattle, Washington; Departments of bRehabilitation Medicine and cPediatrics, University of Washington, Seattle, Washington

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

ABSTRACT

BACKGROUND. Spasticity is a prevalent disabling clinical symptom for children with cerebral palsy. Treatment of spasticity with botulinum toxin in children with cerebral palsy was first reported in 1993. Botulinum toxin provides a focal, controlled muscle weakness with reduction in spasticity. Interpretation of the literature is difficult because of the paucity of reliable measures of spasticity and challenges with measuring meaningful functional changes in children with disabilities.

OBJECTIVE. This study documents the effects of botulinum toxin A injections into the gastrocnemius muscles in children with spastic diplegia. Outcomes are evaluated across all 5 domains of the National Centers for Medical and Rehabilitation Research domains of medical rehabilitation.

METHODS. A randomized, double-masked, placebo-controlled design was applied to 33 children with spastic diplegia with a mean age of 5.5 and Gross Motor Function Classification System Levels of I through III. Participants received either 12 U/kg botulinum toxin A or placebo saline injections to bilateral gastrocnemius muscles. Outcomes were measured at baseline and 3, 8, 12, and 24 weeks after injection.

RESULTS. Significant decreases in the electromyographic representation of spasticity were documented 3 weeks after botulinum toxin A treatment. A significant decrease in viscoelastic aspects of spasticity was present at 8 weeks, and subsequent increases in dorsiflexion range were documented at 12 weeks for the botulinum toxin A group. Improvement was found in performance goals at 12 weeks and in maximum voluntary torque and gross motor function at 24 weeks for the botulinum toxin A. There were no significant differences between groups in satisfaction with performance goals, energy expenditure, Ashworth scores, or frequency of adverse effects.

CONCLUSIONS. The safety profile of 12 U/kg of botulinum toxin A is excellent. Although physiologic and mechanical effects of treatment with botulinum toxin A were documented with functional improvement at 6 months, family satisfaction with outcomes were no different. Communication is needed to ensure realistic expectations of treatment.
Since the first report on the use of botulinum toxin (BTX) to treat spasticity in children with cerebral palsy (CP) was published in 1993, there have been >100 articles addressing the intervention. The present literature, including several meta-analyses, suggests that BTX provides focal, controlled muscle weakness and, by implication, reduction in spasticity. The paucity of reliable measures of spasticity and the difficulty in measuring meaningful changes in function in children with disabilities makes interpretation difficult.

CP, a common chronic disabling condition of childhood, occurs in 1.5 of 1000 to 3 of 1000 live births, with a similar prevalence rate into adolescence and adulthood. Spasticity is a prevalent disabling clinical symptom seen in persons with CP. Spastic diplegia is a very common form of CP, presents with a wide range of ambulatory outcomes, and is most frequently accompanied by ankle joint spasticity. To justify an intervention, there must be a purpose for treating spasticity that is meaningful to patients and their caregivers. Goals of treatment may include simple aims such as reducing pain, increasing range of motion (ROM), preventing secondary medical complications such as contractures or skin breakdown, or better tolerance of splinting and casting. Most patients and their families identify improved active function related to walking as the most important outcome. Not all patients identify the same functional goals nor do they attach the same priorities to those goals.

We comprehensively studied botulinum toxin type A (BTX-A) injections to gastrocnemius muscle in children with CP by using 14 outcome measures distributed among the 5 domains of science relevant to rehabilitation medicine as defined by the National Centers for Medical and Rehabilitation Research (NCMRR) of the National Institutes of Health. These domains include pathophysiology, impairment, functional limitation/activity, disability/participation, and societal limitation/contextual factors (see Tables 1 and 2). This approach afforded a 360° view of the intervention from the pathophysiologic to the societal level. For this project, we combined functional limitation with disability.

We proposed that BTX-A injection into plantar flexor muscles in children with spastic diplegia would produce a reduction of spasticity and a change in function that could be measured across all 5 domains. This project addressed 8 clinical questions: (1) What is the natural history of the weakness associated with change in neuromuscular junction transmission after injection with BTX? (2) Does injection with BTX-A result in clinically relevant reductions in spasticity, measured by changes in torque and stiffness across the ankle joint, when compared with similar untreated patients? (3) What is the magnitude and durability of reduction of spasticity after treatment with BTX-A? (4) Does injection with BTX-A increase mechanical efficiency, as demonstrated by improvement in the Gross Motor Function Measure (GMFM), reduction of energy expenditure, and improvement of gait? (5) What changes in motor function...
typically occur during the 6 months after injection with BTX-A? (6) Does injection with BTX-A result in a meaningful reduction in disability as described by the pediatric patient and his/her parents? (7) Does injection with BTX-A reduce societal limitation in the context of community and school activities? (8) Are there serious adverse events or complications of injection with BTX-A during the first 6 months after the procedure?

Our primary hypothesis was that injection with BTX-A would reduce spasticity in plantar flexor muscles of children with spastic diplegia for 6 months’ duration and that we could reliably measure that reduction in spasticity. We predicted that the pathophysiologic improvements and reductions in spasticity would return to baseline levels at 6 months. We also hypothesized that injections would improve functional ability and societal participation that although temporary, would be significant and meaningful.

METHODS

Thirty-three children were assigned randomly to either an experimental group (injection with 12 U/kg BTX-A) or a comparison group (sham injection of normal saline) by sealed envelope randomization code generated by a study statistician and pharmacist. Double-masked assessments were made at baselines 1 and 2 (7 days apart) and at 3, 8, 12, and 24 weeks postinjection (see Table 1). Injections occurred at baseline number 2 after masked testing was completed. The timing of the first postintervention assessment at 3 weeks after injection was chosen on the basis of previous published reports suggesting that the maximal effect of the neuromuscular blockade was obtained at 21 days. The remaining time points were chosen to assess BTX-A effect duration and document its dissipation. At the end of 24 weeks, the randomization code was unmasked and patients in the sham injection group had the opportunity to cross over and have injection with BTX-A.

Study participants underwent conscious sedation with oral midazolam to improve tolerance of injections. Injections were performed solely by the principal investigator using electromyogram guidance at 2 injection sites in each of the medial and lateral heads of the gastrocnemius muscles. A total dosage of 12 U/kg (never exceeding 400 U) was distributed among left and right gastrocnemius muscles. Assessments were performed in the same order on each assessment day following the schedule in Table 1. The active drug in this clinical trial was botulinum toxin type A (Botox; Allergan Pharmaceuticals, Irvine, CA). Each 100 U of BTX-A was diluted with 1 mL of preservative-free 0.9% sodium chloride for injection. The placebo was 0.9% sodium chloride for injection, prepared to the same volume calculated for the active drug. The study drug was prepared by using the aseptic technique on the day of dosing. All doses were mixed within 4 hours of the time of the injection. The Investigational Drug Service at Children’s Hospital and Regional Medical Center (CHRMC) performed the randomization and prepared all doses of the study drug.

Participants were randomly assigned into 4 strata on the basis of age (3–7 vs 8–12 years) and whether they were receiving oral baclofen. Preliminary power analysis was based on the primary outcome measures of spasticity and motor function (ie, spasticity measurement system [SMS] and GMFM). On the basis of the then-current estimates of variance of primary outcome measures, with 18 subjects in each group, we predicted the ability to detect a 7.2-point change in total GMFM percent scores with a power of 0.9 at a 2-tailed significance level of .05. The same sample size would yield a 15.3-point change in path length on the SMS with a power of .8 and will provide a 2-tailed significance of .05. On the basis of our previous experience with measurement of GMFM and SMS, preliminary analysis of the statistical power of this study suggested that a minimum of 36 subjects would provide adequate power to be confident of our results. The investigators, study coordinators, physical therapists, and participants were masked to treatment assignment. Only the staff of the Investigational Drug Service was unmasked. The institutional review board at CHRMC granted approval for the study.

Study Participants

The multidisciplinary Spasticity Management Clinic team at CHRMC initially screened all potential participants. Participants were approached for enrollment if they presented with the clinical indications for gastrocnemius BTX-A treatment and fit the inclusion/exclusion criteria. Inclusion criteria were (1) age 3 to 12 years; (2) diagnosis of spastic diplegia, defined as spastic motor impairment with more involvement of lower limbs than upper, fairly good trunk and head control, and little or no bulbar involvement; (3) community or indoor ambulation status12–14; (4) cooperative and tolerant to testing procedures during clinic screening; (5) no fixed musculoskeletal deformities >15° (ie, normal knee extension range of motion: 0°, ankle dorsiflexion: +15°), (6) stable social environment; (7) ongoing physical therapy (PT) of a minimum of one 60-minute session per week; and (8) pharmacological treatment (baclofen) for spasticity was not a basis for exclusion. Participants were maintained on a stable, fixed dose 1 month before and throughout the 6-month study period. Orthotic type, wearing schedule, PT regimen, and/or the clinical need for serial casting were documented but not controlled during the study period. All participants received direct PT for a minimum of 1 hour per week during the study period. The therapy regimen was left to the discretion of the treating physical therapist.

A total of 106 children were screened between October 1997 and September 2001, of whom 60 were eligible.
and approached for participation in the trial. Final enrollment was 33 (55% enrollment rate). Informed consent was obtained from all participants and their guardians. Baseline demographics and outcome measures are presented in Table 3. Participants averaged 5.5 years in age (range: 3.1–11.9), and mobility impairments fell into Gross Motor Function Classification System (GMFCS) levels I to III.15,16 There were no significant differences between groups in the baseline measures, except for walking heart rate (P < .003; Table 3). Seventeen participants were randomized to receive BTX-A, and 16 received saline injections. Six children were taking oral baclofen and were equally distributed by treatment group. None of the participants received BTX injections before participation.

Measurements
NCMRR domains of science were represented by 14 unique outcome measures. See Tables 1 and 4 for specific measures by domain. Primary outcome measures for the domains of impairment and functional limitation/disability were total and elastic path length as a measure of spasticity (SMS), and the GMFM (GMFM-88 and GMFM-66). Secondary outcomes included: quantitative electromyographic kinesiology (QEK) for the pathophysiology domain, Ashworth, deep-tendon reflexes (DTR), clonus, ankle dorsiflexion passive range of motion (PROM), and maximum voluntary torque (max torque) of the gastrocnemius muscle for the impairment domain. The functional limitation/disability domains were also measured by energy cost index (ECI) and the performance portion of the Canadian Occupational Performance Measure (COPM). The satisfaction portion of the COPM and Goal Attainment Scaling (GAS) was administered to assess the societal limitation domain.

The QEK and SMS were collected by a study engineer and research assistant masked to group assignment in the research motion analysis laboratory at the University of Washington. Masked research physical therapists and nurses collected the remaining outcome measures in the pediatric clinical research center at CHRMC. A masked research nurse conducted a structured adverse event interview at each research visit in person and by telephone for the 8-week visit. Adverse events reported during those interviews were coded for date of onset and resolution, severity level, relationship to CP and/or treatment, effect on treatment, medical treatments administered, and outcome.

SMS
An electromechanical method of eliciting and measuring spasticity was performed by using automated measurement of ankle stiffness in response to passive ankle movements of various frequencies.17,18 The participant was tested in the prone position and on the leg exhibiting the greatest motor impairment (greatest Ashworth score at screening visit) and with adequate ROM to perform the test. The foot was inserted into a loof bind-

<table>
<thead>
<tr>
<th>TABLE 3</th>
<th>Comparison of Study Sample Baseline Characteristics by Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BTX-A (N = 17)</td>
</tr>
<tr>
<td>Gender, % male</td>
<td>70.59</td>
</tr>
<tr>
<td>Age (SD), y</td>
<td>5.38 (2.06)</td>
</tr>
<tr>
<td>Race, n (%) white</td>
<td>7 (41.18)</td>
</tr>
<tr>
<td>Baseline GMFCS level, %</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>47</td>
</tr>
<tr>
<td>II</td>
<td>41</td>
</tr>
<tr>
<td>III</td>
<td>12</td>
</tr>
<tr>
<td>QEK, MRV, μV</td>
<td>19.57 (17.17)</td>
</tr>
<tr>
<td>SMS total path length, median (SD), N/m/radian</td>
<td>25.3 (76.2)</td>
</tr>
<tr>
<td>SMS elastic path length, median (SD), N/m/radian</td>
<td>21.5 (48.9)</td>
</tr>
<tr>
<td>L/R Ashworth, median score</td>
<td>2.50</td>
</tr>
<tr>
<td>L/R DTR, median score</td>
<td>3.00</td>
</tr>
<tr>
<td>L/R ROM, median degrees</td>
<td>5.00</td>
</tr>
<tr>
<td>L/R Beats clonus, median</td>
<td>5.50</td>
</tr>
<tr>
<td>QEK max torque (mean), N/m</td>
<td>9.92 (10.63)</td>
</tr>
<tr>
<td>GMFM-88 (median total % score)</td>
<td>71.0 (82.5)</td>
</tr>
<tr>
<td>GMFM-66 (median % score)</td>
<td>67.8 (43.4)</td>
</tr>
<tr>
<td>ECI walking heart rate (mean), beats per min</td>
<td>137.5 (14.7)</td>
</tr>
<tr>
<td>ECI walking speed (mean), m/min</td>
<td>37 (16.27)</td>
</tr>
<tr>
<td>ECI (mean walking heart rate/walking speed)</td>
<td>4.94 (3.16)</td>
</tr>
<tr>
<td>COPM performance (mean)</td>
<td>3.75 (1.27)</td>
</tr>
<tr>
<td>COPM satisfaction (mean)</td>
<td>4.03 (1.59)</td>
</tr>
<tr>
<td>GAS (mean)</td>
<td>48.18 (00)</td>
</tr>
</tbody>
</table>

L/R indicates left/right.

a χ²;
b 2-tailed t test;
c Mann-Whitney U.
The GMFM was used to assess changes in gross motor skill and mobility. The GMFM is a criterion reference tool designed to measure change in gross motor function over time in children with motor impairment, and has been validated for sensitivity to change in children with CP. Only items from the standing and walk/run/jump dimensions were administered. Both the GMFM-88 and GMFM-66 scores were used to take advantage of the improved scaling with the GMFM-66. The masked research physical therapists achieved interrater reliability at >80% point-by-point agreement and maintained this for the study’s duration by scoring of standard videotaped GMFM evaluations every 6 months. The 2 research physical therapists had a minimum of 8 years clinical experience each in the administration, use, and scoring of the GMFM-88.

### QEK and Max Torque

The QEK consisted of training the participants to maximally contract the gastrocnemius muscle group and then measuring the electromyograph-generated voltage amplitude and torque responses. The QEK evaluation was obtained in conjunction with the SMS evaluation. To perform this test, the SMS was used in a static (isometric) configuration with the foot positioned in relative plantar-flexion, that is, at the “low” end of the 5° movement range applied for the SMS measurement. The subject was provided with visual feedback of the torque response to encourage the maximal plantar-flexion effort. Data were collected over a 10-second period using the same hardware as described above for the SMS, but sampled at 500 Hz with electromyograph band-pass filtering over a range of 16 to 160 Hz. The peak plantar-flexing torque over the trial period was determined.

### TABLE 4 Outcome Measure Change From Baseline According to NCMRR Domains and Treatment Group at 3-, 8-, 12-, and 24-Week Assessments

<table>
<thead>
<tr>
<th>NCMRR Domains</th>
<th>3 wk</th>
<th>8 wk</th>
<th>12 wk</th>
<th>24 wk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean QEK (µV)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median L/R DTR</td>
<td>1 (3)</td>
<td>0 (3.5)</td>
<td>.005</td>
<td></td>
</tr>
<tr>
<td>Median L/R ankle ROM</td>
<td>-7.5 (25)</td>
<td>-5.0 (27)</td>
<td>.76</td>
<td></td>
</tr>
<tr>
<td>Median L/R beats clonus</td>
<td>-3 (15)</td>
<td>0 (15)</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>Mean max torque (N·m)</td>
<td>-0.88 (35.8)</td>
<td>-1.9 (8.6)</td>
<td>.77</td>
<td></td>
</tr>
</tbody>
</table>

1/2 L/R indicates left/right; WHR, walking heart rate (beats per minute); WS, walking speed (m/minute).

1 Unpaired t test.

2 Wilcoxon Mann-Whitney U.

ing aligned to the ankle joint and positioned at the subject’s maximum dorsiflexion. The foot was passively rotated over 5° at frequencies of 3 to 12 Hz via a custom-designed, computer-controlled actuator. Ankle torque and displacement signals were sampled at 2048 Hz via laptop computer (Apple PowerBook, Cupertino, CA) running custom software written in LabVIEW (National Instruments DAQCard 1200). Torque values were recorded in Newton-meters (N·m) and the path length in Newton-meters per radian. Data were collected over the course of 30 randomly applied 20-second trials (3 trials at each of 10 frequencies). Concomitant surface electromyogram measurements of tibialis anterior and gastrocnemius muscles, using a TE4 electromyograph (TECA Corp, White Plains, NY), provided visualization of the reflex responses and feedback to the examiner to ensure the subject was otherwise relaxed. Inertial and drag torques were computed and deducted from the total, resulting in computation of net passive resistance and spasticity-induced elastic (ie, position-dependent) and viscous (velocity-dependent) ankle stiffness. Variation of ankle stiffness over the range of applied frequencies formed the basis for quantifying spasticity as a total path length and an elastic path length parameter. Larger path length values indicated greater variations in stiffness with frequency that, in turn, reflected an increased reflex mediated or spastic response.19

### GMFM

The GMFM was used to assess changes in gross motor skill and mobility. The GMFM is a criterion reference tool designed to measure change in gross motor function over time in children with motor impairment, and has been validated for sensitivity to change in children with CP. Only items from the standing and walk/run/jump dimensions were administered. Both the GMFM-88 and GMFM-66 scores were used to take advantage of the improved scaling with the GMFM-66. The masked research physical therapists achieved interrater reliability at >80% point-by-point agreement and maintained this for the study’s duration by scoring of standard videotaped GMFM evaluations every 6 months. The 2 research physical therapists had a minimum of 8 years clinical experience each in the administration, use, and scoring of the GMFM-88.
along with surface electromyographic signal parameters by using custom software written in LabVIEW. A running average of the torque over a 1-second window was computed and the electromyographic signal in this averaging window, during which the torque was maximized, was extracted. This represented the electromyographic activity, which occurred at the time of highest mean plantar-flexing torque. The extracted electromyographic signal was rectified and its mean value calculated and reported as the mean rectified voltage (MRV). MRV is a neurophysiologic representation of the total number of motor units available for muscular contraction. The change in MRV is proportional to the magnitude of the neuromuscular blockade. A brief trial period was conducted to train the participants in applying maximum plantar-flexing torque. With the aid of the visual torque feedback, family members and test personnel provided verbal encouragement to the subject to maximize the level of plantar-flexion torque generated. Three QEK determinations were obtained, and the trial exhibiting the highest peak torque response was used because it reflected the most effective attempt.

**Physical Examination Measures**

Masked research physical therapists collected modified Ashworth scores on the gastrocnemius muscle group, PROM, DTRs, and clonus of the ankle in a standardized manner. Interrater reliability of the research physical therapists was established for the Ashworth scale before the study. The research physical therapists had >28 years combined pediatric clinical experience. The principal investigator (Dr Hays) trained the research therapists in the administration and scoring of ankle DTRs and clonus. DTRs were coded from 0 to 4 on an ordinal scale. Clonus was coded by beats palpated.

**Energy Cost Index**

To evaluate the effect of treatment on energy expenditure, the ECI was measured by the research therapists and nurses. Participants walked at a self-selected speed along an oblong 20-meter track for 5 minutes. The participant’s average heart rate and walking speed were recorded. ECI was calculated as the average number of heart beats per unit distance walked.

**COPM and GAS**

Using a semistructured interview recommended by the COPM guidelines, 3 performance outcomes (goals) were identified at baseline by the parent and child with the research physical therapist. These included such goals as “decrease in falling on the playground at school or climb stairs without physical help.” Serial measurements of the COPM performance and satisfaction scores were ranked by importance and then measured by report from the parent. The same 3 performance outcomes (goals) were then structured for goal attainment scaling in the GAS through semistructured interview and by collaboration between the research physical therapist, parent, and child. The formula derived by Kiresuk and Sherman was used to calculate standardized scores (t score), reflecting the composite change for the 3 a priori goals over the 6-month study period.

**Adverse Events**

Structured adverse event interviews were administered in person by the research nurse at all visits except the 8-week visit, which was done by telephone. All events reported during those interviews were coded for severity and relationship to CP and BTX-A or saline injection.

**Analysis**

The data were entered into the SPSS for Windows 10.05 (SPSS Inc, Chicago, IL) software format for analysis. Change scores were calculated from the average of the 2 baseline assessments to each follow-up assessments time point for the primary outcomes (Table 1). At each follow-up time point, the differences in change scores were compared between the BTX-A and placebo groups with the Wilcoxon Mann-Whitney U nonparametric test for the primary outcome variables. Secondary outcome measures of Ashworth scale, clonus, and DTR change scores at each follow-up time point were also assessed by the same nonparametric statistic because of the lack of a normal distribution. Change score from baseline to each follow-up time point for QEK MRV, max torque, COPM performance, and satisfaction, GAS, and ROM were compared between treatment groups by an unpaired t test. The frequency of adverse events by group was analyzed by χ². Baseline walking heart rate was found to be significantly different between treatment groups by an unpaired t test. The frequency of adverse events by group was analyzed by χ². Baseline walking heart rate was found to be significantly different between treatment groups (Table 3) and was corrected for in the regression analysis of ECI change from baseline. Change-score statistics (mean, SD or median, range) for the appropriate tests for each outcome measure are displayed in Table 4.

**RESULTS**

**Pathophysiology**

The known neurophysiologic effect of BTX injections compared with placebo was confirmed at 3 weeks’ postinjection by a significant decrease in QEK (MRV-μV, P < .05). This neurophysiological difference was no longer present by 8 weeks. By 24 weeks, the BTX group QEK had surpassed the placebo group, although not significantly (Fig 1 and Table 4).

**Impairment**

Significant decreases in SMS total (P < .04) and elastic path length (P < .05) changes from baseline were documented at 8 weeks’ postinjection for the BTX group compared with the placebo group. A change in path length for the BTX group compared with the placebo
from baseline was present at 12 weeks but did not reach significance (see Figs 2 and 3 and Table 4). No significant differences in change scores by group were found for Ashworth change scores at any follow-up time point (see Table 4). The changes in the clinical measures of Achilles DTR and clonus were significantly decreased in the BTX group at 3 weeks only (Table 4). Ankle dorsiflexion PROM with the knee extended significantly increased for the BTX group ($P < .05$) compared with the placebo group at 12 weeks with a mean difference of 4.2 degrees (Fig 4 and Table 4). QEK maximum torque changes from baseline were significantly greater for the BTX group at 24 weeks postinjection ($P < .03$) (Fig 5 and Table 4).

Functional Activity/Disability

GMFM-88 total scores approached a significant difference at 3 weeks, with the BTX group scoring higher. Both groups improved their change from baseline scores over the study period with a significant median difference documented at 24 weeks ($P < .001$; Fig 6 and Table 4). GMFM-66 change from baseline scores for the BTX group was also significantly greater than the placebo group at 24 weeks (3.1 BTX vs 1.2 control, $P < .03$; Fig 7 and Table 4). No significant differences between treatment group change scores were found at any follow-up time point for ECI (Table 4). COPM performance change scores from baseline were significantly greater for the
BTX group at 12 weeks (1.7 BTX vs 1.2 placebo; \(P < .04\)). At 24 weeks, the BTX group continued with higher change scores than the placebo group but did not reach significance (Fig 8 and Table 4). The significant changes from baseline documented for individuals in the BTX group suggest improved gait and related upright motor skills (ie, stair climbing, 1-leg balance, jumping).

**Societal Participation**
The COPM satisfaction change scores and GAS t scores increased from baseline to 12- and 24-week follow-up for both groups. No significant differences in change scores between groups were found at 12 or 24 weeks’ follow-up (\(P > .12–.98\)) in Figs 9 and 10 and Table 4.

**Adverse Events**
A total of 56 adverse events potentially having any relationship to treatment (injection of BTX-A or saline) were reported during the 6-month study period for both treatment groups. The frequency of adverse events by treatment group (30 for BTX, 26 for placebo) was not significantly different between the groups (\(P = .22\)). Six of these events required ibuprofen for muscle soreness at injection site (3 per treatment group), and 3 decreased their activity level for 24 hours after injection.

**DISCUSSION**
This is the first randomized, controlled trial of BTX treatment to report outcomes globally across all domains of
the NCMRR framework. We have attempted to address a question that has been characterized by Sheean as the "holy grail" of BTX investigation, which is to document not only physiologic effect but functional benefit. We have also attempted to avoid the 4 most common barriers to good clinical research regarding BTX. These include inappropriate outcome measures, inappropriate patient selection, inappropriate injection protocols, and poor general study design.

We were gratified that our physiologic measure of motor strength reduction after treatment with BTX-A, the MRV measure of the QEK, differed significantly from the placebo group and mirrored the descriptions of BTX effect described in experimental animal studies. This finding, a significant reduction at 3 weeks, indicated that the neuromuscular junction transmission was genuinely altered and that the injection technique was valid and reliable.

Significant reduction in spasticity was documented. The most robust evidence was found in the change in mean total and elastic path lengths measured at 8 weeks by using the SMS. This finding was supported by significant changes in the DTRs and clonus at the ankle joint. The increases in dorsiflexion ROM at 12 weeks may have been related to the changes in the viscoelastic properties of the muscle that followed the initial neuromuscular blockade and were maintained through more effective active use of the muscle in gait. Significant changes in the voluntary torque documented at 24 weeks may most likely be a direct response to the improved ROM. This apparent chain of improvements in different aspects of function is consistent with real physiologic change.

GMFM and gastrocnemius torque results show significant consistent changes in mechanical efficiency and motor function at 6 months after treatment. This was surprising and may represent the time required to adjust to the changes in motor strength, ROM, and the reduction of interfering spasticity. The later changes in motor function and strength support an argument for consistent therapy to maintain and maximize the effects of BTX-A treatment, to facilitate acquisition of new skills or improved efficiency, and a longer period before reinjection with BTX. No significant change in energy expenditure using the ECI was found.

We measured a significant change in the experimental subjects’ performance with the COPM after treatment compared with the placebo group, but we were unable to demonstrate a meaningful difference in either the satisfaction scores of the COPM or the mean change in t scores via GAS. This result was intriguing.

The results of this project both parallel and contrast with other randomized, control trials of BTX in ambulatory youth with CP. Sutherland and colleagues documented improved ankle dorsiflexion and swing during walking with 3 dimensional gait analysis at 3 months’ post-BTX injections. The improvement in GMFM-88 total and GMFM-66 scores found in this project at 6 months reflects global changes in walking skill and are in line with the gait analysis outcomes found by Sutherland. Using a clinical observational scale to quantify parameters of gait, Koman et al also documented improved gait patterns at 12 weeks’ postinjection with BTX in a mixed sample of youth with hemiplegia and diplegia CP. Komen and colleagues also noted significant improvements in ankle range of motion which is consistent with data we report here in Table 4. In 2000, Ubhi and colleagues reported significant improvements in GMFM walking skills at 3 months after BTX injections for ankle equinus in a group of children with hemiplegia (30%) and diplegia motor involvement. We did not find this improvement in gross motor skills in our study sample until 6 months after injection. The lack of significant improvement in the physiologic cost index reported by Ubhi and coauthors is consistent with our findings for ECI. Their results contrast with the changes documented in ankle ROM (Table 4). These conflicting results may be a function of mixed study samples containing both youth with hemiplegia and diplegia as well as older children studied. None of these studies looked at outcomes at 6 months after injection.

Reddihough et al followed subjects for 6 months and documented no significant improvement in GMFM scores at 3 and 6 months after BTX injections in children with spastic diplegic CP. The age of their study population was similar to subjects in our project, but >60% of their study sample were GMFCS levels III and IV, suggesting they were treating subjects with lower walking skills than the subjects in our study sample where 50% were GMFCS level II. Most recently, Scholtes and colleagues in the Netherlands also reported improved GMFM scores that were similar to our results at 6 months. They used multilevel BTX-A injections and provided comprehensive rehabilitation in 46 children across GMFCS levels I-IV.

CONCLUSIONS
It seems that the physiologic and mechanical effects of treatment with BTX-A are genuine and measurable in youth with spastic diplegia CP. However, these effects may not create enough change in the patients’ function or the families’ perception of function to register as a meaningful improvement in their societal participation. It is possible that these changes were too subtle to be recognized with conventional satisfaction measures. The failure to demonstrate a match between the measured effect and the perceived benefit of treatment is perhaps 1 of the most important findings of our study. Detailed information describing expected functional improvements after treatment, based on studies such as this one, must be communicated to families in useful, understandable terms to ensure that they enter into treatment
with truly informed consent. Our study suggests that patients and families must have realistic expectations about BTX treatment before they enter into it or risk experiencing measurable effect without perceived benefit.

ACKNOWLEDGMENTS

Funding for this project was through the NCMRR, National Institutes of Health grant RO1 HD35750, and grant M01-RR-00037 of the University of Washington General Clinic Research Center from the National Center for Research Resources, National Institutes of Health. The BTX-A was supplied by Allergan Pharmaceuticals, Irvine, CA.

REFERENCES

Early Intervention and Recovery Among Children With Failure to Thrive: Follow-up at Age 8

Maureen M. Black, PhD\textsuperscript{a}, Howard Dubowitz, MS, MD\textsuperscript{b}, Ambika Krishnakumar, PhD\textsuperscript{b}, Raymond H. Starr, Jr, PhD\textsuperscript{c}

\textsuperscript{a}Department of Pediatrics, University of Maryland School of Medicine, Baltimore, Maryland; \textsuperscript{b}Department of Child and Family Studies, Syracuse University, Syracuse, New York; \textsuperscript{c}Department of Psychology, University of Maryland Baltimore County, Baltimore, Maryland

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

ABSTRACT

OBJECTIVES. We sought to examine the impact of a randomized, controlled trial of home visiting among infants with failure to thrive on growth, academic/cognitive performance, and home/classroom behavior at age 8.

METHODS. Infants with failure to thrive (N = 130) or adequate growth (N = 119) were recruited from pediatric primary care clinics serving low-income, urban communities. Eligibility criteria included age <25 months, gestational age >36 weeks, birth weight >2500 g, and no significant medical conditions. Evaluation included anthropometrics, Bayley scales, maternal anthropometrics, demographics, negative affect, IQ, and the Home Observation for Measurement of the Environment scale. Infants with failure to thrive were treated in an interdisciplinary growth and nutrition clinic and randomized into clinical-intervention-plus-home-intervention or clinical-care-only groups. The home-visiting curriculum promoted maternal sensitivity, parent-infant relationships, and child development. Follow-up visits were conducted by evaluators who were unaware of the children’s growth or intervention history. At age 8, the evaluation included anthropometrics, the Wechsler Intelligence Scale for Children III, and the Wide Range Achievement Test, Revised. Mothers completed the Child Behavior Checklist and teachers completed the Teacher Report Form.

ANALYSIS. Multivariate analyses of variance were used to examine differences in growth, cognitive/academic performance, and home/school behavior, adjusted by maternal education, public assistance, and, when appropriate, infant Bayley score, maternal BMI, height, negative affect, IQ, and Home Observation for Measurement of the Environment scores.

RESULTS. Retention was 74\% to 78\%. Children in the adequate-growth group were significantly taller, heavier, and had better arithmetic scores than the clinical-intervention-only group, with the clinical-intervention-plus-home-intervention group intermediate. There were no group differences in IQ, reading, or mother-reported behavior problems. Children in the clinical-intervention-plus-home-intervention group had fewer teacher-reported internalizing problems and better work habits than the clinical-intervention-only group.

CONCLUSIONS. Early failure to thrive increased children’s vulnerability to short stature, poor arithmetic performance, and poor work habits. Home visiting attenuated some of the negative effects of early failure to thrive, possibly by promoting maternal sensitivity and helping children build strong work habits that enabled them to benefit from school. Findings provide evidence for early intervention programs for vulnerable infants.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-1657
doi:10.1542/peds.2006-1657

Key Words
failure to thrive, early intervention, home visiting, longitudinal follow-up

Abbreviations
FTT—failure to thrive
AG—adequate growth
FTT-HI—clinical intervention plus home intervention
FTT-CO—clinical intervention only
MDI—Mental Development Index
PDI—Psychomotor Development Index
HOME—Home Observation for Measurement of the Environment
AFDC—Aid to Families With Dependent Children

Accepted for publication Mar 1, 2007

Address correspondence to Maureen Black, PhD, Department of Pediatrics, University of Maryland School of Medicine, 737 W Lombard St, Room 161, Baltimore, MD 21201. E-mail: mblack@peds.umaryland.edu
During the first years of life when energy needs are high, growth serves as an objective measure of children’s well-being. Failure to thrive (FTT) occurs when infants’ rate of weight gain is below expectations based on age- and gender-specific growth charts. Until recently, many investigators relied on hospitalized or referred samples of children with FTT and found that FTT was associated with long-term deficits in height, weight, cognitive and academic performance, and behavior. However, hospitalized and referred patients are likely to represent the most extreme and complex cases of FTT. Most children with FTT are treated as outpatients, consistent with recommendations from the American Academy of Pediatrics’ Committee on Nutrition and managed in specialized, interdisciplinary clinics.

Investigators relying on population- and community-based samples have found that by school age, most children with a history of FTT have experienced growth recovery. Although many continue to be shorter than age-matched peers, they rarely experience growth deficits indicative of severe malnutrition. When cognitive and academic performance have been considered, by school age, children with FTT recruited from primary care or community sites achieved IQ scores that were ~4.2 points lower than children with a history of adequate growth (AG). These findings suggest that early FTT may have a small, though potentially important, impact on children’s cognitive performance.

Although a recent meta-analysis of 60 home-visiting programs introduced early in life showed encouraging findings on children’s cognitive and social-emotional development, the 6 home visitation trials conducted among children with FTT have reported inconsistent effects on children’s growth and development. Haynes and colleagues studied 50 hospitalized children with FTT and 26 comparison children and found no effect of short-term home visitation on growth, development, or parent-child interaction patterns. Drotar and Sturm used a randomization procedure to assign children into 1 of 3 home intervention programs and found no difference in growth or cognitive performance at 3 years of age. Casey and colleagues reported that among 914 preterm children with low birth weight, the incidence of FTT did not differ on the basis of participation in an early intervention program. However, among the children with FTT, there were beneficial effects of the intervention on cognition and behavior at age 3. Wright and colleagues studied 229 children with FTT identified through community screening and showed that children assigned to the intervention group who received home health visitors were heavier and taller at 3 years old than children assigned to the control group. Raynor and colleagues enrolled 83 children with FTT and provided home visiting to 42 of them. After 12 months, both groups experienced a significant increase in weight, with no differential effects of home visiting. We conducted a weekly home-based intervention among 130 infants and toddlers with FTT recruited from primary care. After 1 year, there were no differences in height or weight, but infants in the intervention group had better cognitive performance and caregivers who were more responsive and child focused than infants in the control group. Two years after the home intervention ended, when the children were 4 years old, children in the intervention group had better cognitive scores and were more socially interactive than children in the control group, but only if their mothers did not report negative affect.

To examine whether early intervention altered the children’s developmental course, we followed children with FTT through their school-age years, along with a cohort of adequately growing children from the same low-income communities. The primary objective of the follow-up was to examine the long-term impact of home-based intervention on children’s growth, academic and cognitive performance, and home and school behavior at age 8.

**METHODS**

**Participants**

Infants were recruited from 1989–1992 from pediatric primary care clinics that serve low-income, urban communities. Eligibility criteria included age <25 months, gestational age >36 weeks, birth weight ≥2500 g, and no congenital problems, disabilities, or chronic illnesses.

Children in the FTT group had to meet 1 of 2 criteria using age- and gender-specific National Center for Health Statistic growth charts: sustained weight for age <5th percentile or weight for length <10th percentile. Children in the AG group had to meet 2 criteria: sustained weight for age and weight for length >10th percentile. The groups were matched by gender, race, and socioeconomic status, defined by marital status and dependence on public assistance. All infants were examined by a pediatrician, who also reviewed their medical charts to ensure they met criteria for group classification and there were no known syndromes or obvious major organ system dysfunctions, such as congenital heart disease, to account for the growth failure of the infants in the FTT group.

**Procedure**

Research assistants invited caregivers to participate in a longitudinal research project, using consent procedures approved by the institutional review board of the University of Maryland. More than 90% of eligible caregivers agreed and participated in an initial evaluation that included measures of growth, standardized developmental assessments, and a 60-minute interview of questionnaires on demographics, children’s behavior, and mater-
nal and family functioning. Developmental assessments were administered by psychology graduate students, supervised by a pediatric psychologist. A home visit was scheduled within 2 weeks of the initial evaluation.

Children with FTT were treated in an interdisciplinary clinic. Based on a randomization procedure, stratified by race, gender, and infant age and designed by a statistician to ensure equivalence across groups, children with FTT were randomized to receive either the clinical intervention plus home intervention (FTT-HI) or the clinical intervention only (FTT-CO) (Fig 1). The AG children received standard pediatric primary care.

The children and caregivers returned for evaluations at 4, 6, and 8 years of age. Evaluators were unaware of their growth or intervention status. Caregivers provided the name of the children’s school and requests were sent for information on classroom behavior. Families and teachers were compensated for participating. We report here on the 8-year evaluation.

**Intervention**

Details regarding the home intervention have been published previously and will be summarized. The intervention was based on ecological theory and included a therapeutic alliance between the interventionist and the caregiver; support to the caregiver’s personal, family, and environmental needs; opportunities to model and promote responsive parent-infant interaction; and problem-solving strategies regarding personal, parenting, and children’s issues. The Hawaii Early Learning Program was used as a curriculum guide.

The intervention was delivered by 3, part-time lay home visitors employed by a community-based agency specializing in early intervention. The home visitors received an 8-session training program and were supervised by a community health nurse. The home visitors had portable mats and toys to demonstrate developmentally appropriate activities and to facilitate parent-child interaction. They did not focus on nutrition or feeding behavior and they did not weigh the children. One-hour visits were scheduled weekly for 1 year; the number of visits varied from 0 to 47 (mean: 19.2; SD: 11.5; median: 25; interquartile range: 18–30).

**Baseline Measures**

**Infant Growth**

Children were weighed and measured by a nurse using scales calibrated regularly. Length was converted into age- and gender-specific SD units (z scores) of length for age, adjusted for midparental stature. Weight was converted into age- and gender-specific SD units (z scores) for weight for age and weight for length.

**Infant Mental and Motor Development**

Mental and motor development were assessed by the Bayley Scale of Infant Development. Raw scores for the mental and motor scales were converted to age-normed Mental Development Index (MDI) and Psychomotor Development Index (PDI) scores with a population mean of 100 and an SD of 15.

Maternal negative affect was measured with the Brief Symptom Inventory, a symptom scale in which respondents report on the frequency of symptoms over the preceding 7 days using a 4-point scale, ranging from not at all (1) to extremely (4). Low scores are optimal. Scores from the depression, anxiety, and hostility subscales were combined into a measure of negative affect, with an internal consistency, calculated by Cronbach’s α of .92.

Maternal IQ was measured with the comprehension and vocabulary subscales from the Wechsler Adult Intelligence Scale. The subscales correlated highly with the full-scale IQ.

The home environment was measured by using the Home Observation for Measurement of the Environ-

---

**FIGURE 1**

Description of sample recruitment and retention.
ment (HOME), an observation scale that has been widely used in child development research and has a strong relationship with subsequent intellectual development and achievement. It consists of 45 items organized into 6 scales (emotional and verbal responsivity, avoidance of restriction and punishment, organization of the physical and temporal environment, provision of appropriate play materials, maternal involvement with the child, and opportunities for variety in daily stimulation). Observers were trained until they reached >90% agreement on 10 observations. Reliability was maintained at >90% agreement.

**Age 8 Measures**

At age 8, family demographics and maternal negative affect were repeated.

**Growth**

Children were weighed and measured by a trained nurse on scales calibrated regularly. Height and weight were converted into age- and gender-specific SD units (z scores) of height for age and weight for age. BMI was calculated by using the formula: weight (kg)/length (m²) and converted to age- and gender-specific z scores on the basis of the 2000 Centers for Disease Control and Prevention tables (ww.cdc.gov/growthcharts).

**Cognitive Performance**

The vocabulary and block-design subtests of the Wechsler Intelligence Scale for Children-III were administered. A composite score was transformed into a standardized score with a mean of 100 and an SD of 15.

**Academic Performance**

The Arithmetic and Reading Subtests of the Wide Range Achievement Test-Revised were administered.

**Home Behavior**

Mothers completed the Child Behavior Checklist, which consists of 118 behavior-problem items rated on a 3-point scale ranging from 0 (not true) to 3 (very true). T scores were obtained for externalizing (eg, delinquency and aggression) and internalizing problem behaviors (eg, depression, anxiety/withdrawal). High scores represent more problems.

**School Behavior**

Teachers completed the Teacher Report Form, consisting of 118 behavior-problem items rated on a 3-point scale ranging from 0 (not true) to 3 (very true). T scores were obtained for externalizing and internalizing problem behaviors, with high scores representing more problems. The Teacher Report Form includes 4 positive classroom work habits (works hard, behaves appropriately, learns, and happy), based on 7-point Likert scales comparing the target child to typical students, ranging from much less (1) to much more (7). Cronbach’s α value for our sample was .87. Teachers were unaware of the children’s growth or intervention history.

**Sample Demographics**

The baseline demographic characteristics of the 189 families who completed the 8-year evaluation indicated that most families received public assistance (Aid To Families With Dependent Children [AFDC] or Medical Assistance: 76.2%) (Table 1). Approximately half of the mothers had not completed high school (54%), and most were unemployed (87%) and single (89%). At baseline, the children with a history of FTT were slightly younger than AG children. With the exception of maternal education, the families of children with and without a history of FTT were similar in most demographic characteristics. There were no group differences in maternal IQ, maternal negative affect, maternal anthropometry, or the home environment among the children in the FTT-HI, FTT-CO, and AG groups.

The children with FTT were shorter and thinner than children in the AG group, as expected. There were no differences in anthropometry between the 2 FTT groups. Children in the FTT groups had significantly lower scores on psychomotor development than those in the AG group; there were no group differences on mental development.

**Analysis Plan**

In keeping with the guidelines for a randomized trial, the analysis was based on assignment of the FTT group into intervention or control, regardless of the amount of intervention received. To examine the effects of early growth failure on children’s growth, cognitive/academic performance, and home and school behavior at age 8, we used multivariate analysis of variance. The analyses examined the effects of group (FTT-HI, FTT-CO, and AG), followed by pairwise comparisons. In the initial analyses, we did not adjust for confounders because our goal was to examine the children’s functioning at age 8, regardless of their early history. In subsequent analyses, we adjusted for confounders.

To identify potential confounders, we examined the relation between baseline indicators of poverty (public assistance and household size), maternal anthropometry (weight and height), and maternal functioning (maternal education, IQ, marital status, negative affect, and the HOME inventory) with children’s growth and functioning at age 8 (Table 2). Receipt of public assistance (AFDC or Medical Assistance) was associated with lower reading scores and worse school behavior. Maternal height and BMI were positively associated with children’s corresponding measures at age 8. Maternal education was positively associated with BMI, height, IQ, reading and arithmetic scores, and classroom behavior, and negatively associated with externalizing behavior at home.
Maternal IQ was positively associated with IQ and reading. HOME scores were positively associated with IQ, reading, arithmetic, and learning in school. Maternal negative affect was positively associated with internalizing and externalizing behavior at home.

We examined associations between children’s baseline anthropometry and performance on the Bayley Scales with their growth and functioning at age 8. Baseline weight for height was positively associated with BMI ($r = 0.33; P < .001$) and baseline length for height was positively associated with BMI ($r = 0.53; P < .001$), and both were associated with reading and arithmetic at age 8 ($r = 0.24–0.33; P < .001$). In addition, Bayley MDI was associated with IQ ($r = 0.27; P < .001$).

To ensure that findings were not confounded by household resources, we controlled for receipt of public assistance and maternal education in all analyses. When analyzing age 8 BMI and height for age, we also controlled for maternal BMI and height for age in the analyses for cognitive/academic performance, we controlled for maternal IQ and HOME. In the analysis for home behavior, we controlled for maternal negative affect. We used estimated marginal means to examine group differences (FTT-HI, FTT-CO, and AG).

To examine how contemporaneous environmental measures were related to child growth and functioning, we examined relations between age 8 measures of public assistance, maternal education, marital status, and negative affect and children’s growth and functioning. There were no significant associations between public assistance and marital status with age 8 growth or functioning. Maternal education was significantly positively associated with all the growth and cognitive/academic measures ($r = 0.18–0.26; P < .05$) and significantly negatively associated with externalizing ($r = -0.17; P = .02$), but not internalizing behavior at home. Maternal negative affect was significantly negatively associated with arithmetic ($r = -0.14; P = -0.04$) and positively associated with internalizing and externalizing behavior at home ($r = 0.35–0.45; P < .001$). The analyses were
repeated, controlling for maternal education and negative affect at age 8.

For unadjusted analyses, we set \( \alpha \) at .05. Given the longitudinal nature of the study and the risk of committing a type II error (failing to detect a difference when one exists), we extended \( \alpha \) to .10 for analyses adjusted by covariates.

**RESULTS**

**Retention**

At age 8, there were complete data for 96 (74%) of 130 FTT (47 of 64 HI and 49 of 66 CO) (Fig 1) and 93 (78%) of 119 AG families. With an \( \alpha \) value of .05, this sample size provided power of 0.80 to detect a difference of 0.6 SDs (~8 points) in measures of IQ and academic performance among children in the 3 groups. Demographic comparisons between those who participated at age 8 and those who did not indicated no group differences in rates of retention and no differences on any background variables, including anthropometric status, gender, race, mental and motor development, maternal variables (height, weight, education, IQ, negative affect, marital status), HOME, household size, or intervention status.

**Growth At Age 8**

In unadjusted analyses, children in the AG group were taller and heavier than children in the 2 FTT groups (Table 1). There were no significant differences in height for age \( z \) score between the 2 FTT groups, but there was a significant linear trend \( (F = 11.84; P = .001) \), with the children in the FTT-HI group occupying an intermediate position between children in the FTT-CO and AG groups. Rates of stunting (height for age less than \(-2 \) \( z \) score) were higher for children in the FTT-CO (8.0%) and FTT-HI groups (6.4%) than in the AG group (0%) \((\chi^2 = 7.16; P < .05)\). When analyses were adjusted for maternal education, receipt of public assistance, and maternal anthropometry, paired comparisons revealed significant differences between all 3 groups, with the FTT-CO children having the lowest scores, FTT-HI intermediate, and AG the highest (Table 3).

When BMI was considered, children in the AG group were heavier than children in the FTT-HI and FTT-CO groups. Although there were no significant differences between the 2 FTT groups, there was a significant linear trend \( (F = 8.96; P = .003) \), with children in the FTT-HI group maintaining an intermediate position between children in the FTT-CO and AG groups. Rates of wasting (weight for age less than \(-2 \) \( z \) scores) were higher for children in the FTT-CO group (6%) than for children in the FTT-HI (0%) and the AG groups (0%) \((\chi^2 = 8.53; P < .05)\). When analyses were adjusted for maternal education, public assistance, and anthropometry, the children in the FTT-CO group were significantly thinner than those in the AG group \((P = .003; \text{Table 3}) \) and marginally thinner than those in the FTT-HI group \((P = .065)\). The FTT-HI group did not differ from the AG group.

**Cognitive/Academic Performance at Age 8**

There were no group differences in IQ (Table 3) in either unadjusted or adjusted analyses. Children in all groups obtained IQ scores that were ~1 SD below those in the standardization sample. Rates of mental retardation (IQ

### TABLE 2 Correlations Between Baseline Public Assistance, Maternal Height, Weight, and Functioning With Growth, Cognitive/Academic Performance, and Home and School Behavior at Age 8

<table>
<thead>
<tr>
<th></th>
<th>Public Assistance</th>
<th>Maternal Height</th>
<th>Maternal BMI</th>
<th>Maternal Education</th>
<th>Maternal IQ</th>
<th>HOME</th>
<th>Maternal Negative Affect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>0.04</td>
<td>0.02</td>
<td>0.24&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.20&lt;sup&gt;b&lt;/sup&gt;</td>
<td>−0.08</td>
<td>−0.03</td>
<td>0.08</td>
</tr>
<tr>
<td>Height, ( z ) score</td>
<td>−0.02</td>
<td>0.34&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.02</td>
<td>0.17&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.09</td>
<td>0.10</td>
<td>−0.13</td>
</tr>
<tr>
<td>Cognitive/academic performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IQ</td>
<td>−0.07</td>
<td>0.03</td>
<td>0.08</td>
<td>0.18&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.21&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.28&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.01</td>
</tr>
<tr>
<td>Reading</td>
<td>−0.16&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.04</td>
<td>0.13</td>
<td>0.20&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.16&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.16&lt;sup&gt;c&lt;/sup&gt;</td>
<td>−0.04</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>−0.04</td>
<td>0.10</td>
<td>0.04</td>
<td>0.15&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.14</td>
<td>0.20&lt;sup&gt;b&lt;/sup&gt;</td>
<td>−0.02</td>
</tr>
<tr>
<td>Home behavior</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internalizing behavior</td>
<td>0.13</td>
<td>0.14</td>
<td>−0.02</td>
<td>−0.12</td>
<td>−0.03</td>
<td>−0.07</td>
<td>0.23&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Externalizing behavior</td>
<td>0.14</td>
<td>0.14</td>
<td>−0.06</td>
<td>−0.17&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.01</td>
<td>−0.11</td>
<td>0.25&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>School behavior</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internalizing behavior</td>
<td>0.14</td>
<td>0.02</td>
<td>0.03</td>
<td>−0.09</td>
<td>−0.08</td>
<td>0.01</td>
<td>0.08</td>
</tr>
<tr>
<td>Externalizing behavior</td>
<td>0.17&lt;sup&gt;c&lt;/sup&gt;</td>
<td>−0.05</td>
<td>−0.03</td>
<td>0.07</td>
<td>−0.03</td>
<td>−0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>Works hard</td>
<td>−0.20&lt;sup&gt;c&lt;/sup&gt;</td>
<td>−0.03</td>
<td>0.04</td>
<td>0.17&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.10</td>
<td>0.14</td>
<td>−0.15</td>
</tr>
<tr>
<td>Behaves appropriately</td>
<td>−0.14</td>
<td>−0.01</td>
<td>0.01</td>
<td>−0.01</td>
<td>−0.04</td>
<td>0.09</td>
<td>−0.09</td>
</tr>
<tr>
<td>Learning</td>
<td>−0.19&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.06</td>
<td>0.04</td>
<td>0.20&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.12</td>
<td>0.18&lt;sup&gt;c&lt;/sup&gt;</td>
<td>−0.11</td>
</tr>
<tr>
<td>Happy</td>
<td>−0.28&lt;sup&gt;b&lt;/sup&gt;</td>
<td>−0.04</td>
<td>−0.06</td>
<td>0.09</td>
<td>0.06</td>
<td>0.02</td>
<td>−0.03</td>
</tr>
<tr>
<td>Overall classroom behavior</td>
<td>−0.21&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.02</td>
<td>0.01</td>
<td>0.13</td>
<td>0.10</td>
<td>0.14</td>
<td>−0.13</td>
</tr>
</tbody>
</table>

<sup>a</sup> \( P < .001 \)

<sup>b</sup> \( P < .01 \)

<sup>c</sup> \( P < .05 \)
School Behavior

In unadjusted analyses, teachers reported more internalizing behavior problems for children in the FTT-CO group compared with children in the FTT-HI group ($P < .05$) or the AG group ($P < .05$). There were no differences in internalizing behaviors between the children in the FTT-HI and AG groups, and no differences in externalizing behaviors among the 3 groups. When positive behaviors were considered, children in the FTT-HI group had significantly higher scores on the works-hard and learning subscales than children in the FTT-CO group ($P < .05$).

In adjusted analyses, children in the FTT-CO group had significantly higher scores in internalizing behaviors ($P = .008$) and marginally lower scores in learning ($P = .07$) than children in the AG group. Children in the FTT-HI group had marginally lower scores in internalizing behaviors ($P = .06$), significantly higher scores in works-had ($P = .05$), and marginally higher scores in happy and learning ($P < .10$) than children in the FTT-CO group. There were no differences in school behaviors between the FTT-HI and AG groups (Table 3).

Contemporary Measures at Age 8

Rates of single parenthood (76.8%) and dependence on public assistance (88.9%) continued to be high (Table 1). With the exception of maternal education, there were no group differences in demographics, family characteristics, or maternal negative affect at age 8. The analyses were repeated by using age 8 measures of maternal education, public assistance, and maternal negative affect as covariates, rather than baseline measures. The only difference from findings using baseline covariates is that the difference in arithmetic between children in the FTT-CO and AG groups reached significance ($P = .03$), favoring children in the AG group.

DISCUSSION

Longitudinal Follow-up of FTT

These longitudinal analyses make important contributions to the controversies surrounding the long-term consequences of FTT on children’s growth, cognitive and academic performance, and home and school behavior, and demonstrate the lasting effects of an early home intervention among children with FTT. Our sample avoided referral bias by recruiting children from primary care and yielded almost no baseline differences in family demographics, maternal negative affect, or the home environment, based on the children’s early growth history. Thus, group differences in growth, arithmetic performance, and classroom work habits that occurred at age 8 may be attributed to early growth failure and home intervention.
Growth

By 8 years of age, children with a history of AG were an average of 5.5 cm taller than children with FTT who received an intervention and 6 cm taller than children with FTT who did not. Differences in height remained significant, after controlling for public assistance, maternal height, and maternal education, suggesting that early malnutrition experienced by the children with FTT was severe enough to hinder their linear growth, at least through the school age years. The deficient linear growth among children with a history of FTT is consistent with findings from other follow-up studies.16,41 Early stunting is a marker for chronic undernutrition and has been associated with poor school-age academic performance.42–44

Children with a history of AG were an average of 4 kg heavier than children with FTT who received an intervention and 6 kg heavier than children with FTT who did not, resulting in BMI differences favoring children in the AG group. Previous analyses demonstrated a pattern of recovery whereby the children in both FTT groups experienced catch-up weight gain between recruitment and age 6.27 The findings are similar to growth patterns from developing countries, where linear growth faltering begins early in life and continues through at least preschool years and weight for length faltering occurs in the first 15 months of life, followed by recovery.45 The importance of considering genetic endowment when studying children’s growth is illustrated by the significant associations between mothers’ weight and height with their child’s weight and height at age 8.

Cognition

The lack of significant differences in cognition between children with and without a history of FTT is consistent with recent literature from the United Kingdom, in which children with FTT were recruited from health centers or communities, rather than from hospitals or referral sources.8,16–18,41 In 2 studies, children’s cognitive performance was higher than the scores achieved in the current sample.16,41 For example, in the Boddy study,41 scores on the McCarthy General Cognitive Index were well within the expected range regardless of early growth history (mean: −101.7; SD: −17.8). The differences in cognitive performance between our data and the United Kingdom data may be partially explained by the socioeconomic status of the samples. The United Kingdom samples were population-based and represented a wider range of socioeconomic status than the current sample, which was recruited from primary care clinics serving a very low-income urban community. The United Kingdom samples also differed from the current sample in race (white versus black), maternal age (older versus younger), maternal IQ (higher versus lower) and household composition (2-parent versus single parent). Thus, children in the current sample may have been living in more impoverished households than those in the United Kingdom studies.

The present finding of low cognitive scores among children (~1 SD below the normative sample), regardless of their growth history, is consistent with findings from other samples of low-income children. Evidence from the National Institute of Child Health and Human Development Early Child Care Research Network has shown that children in chronically impoverished families have lower cognitive performance and more behavior problems than children who are not exposed to poverty, partially explained by a lack of stimulating behaviors and home experiences among low-income families.26

A major finding from the current study is evidence that the early caregiving environment plays a critical role in children’s cognitive and academic performance at age 8, regardless of early growth history. On average, the children in our study had cognitive skills slightly below the reference range when they were recruited during infancy. Infant performance on the Bayley was associated with measures of IQ and academic performance at age 8. By age 8, the children’s cognitive and academic performance had declined significantly, possibly because of limited intellectual opportunities in the caregiving environment. Not only did mothers in our sample have IQ scores that were >1 SD below the norm, but maternal IQ and the quality of the HOME, measured when the children were ~15 months of age, were significantly associated with 3 independent assessments of children’s functioning at age 8: a standardized measure of IQ, a standardized test of arithmetic, and teacher assessments of learning skills. These findings suggest that limitations in the early caregiving environment may have hindered children’s cognitive and academic performance regardless of their growth history.

Home Intervention

The protective effects of early intervention are demonstrated by the children’s age 8 growth, academic performance, and school behavior. The home intervention was designed to enhance mother-child relationships and maternal sensitivity by teaching mothers to respond to their child’s bids for interaction. At the conclusion of the intervention, mothers in the intervention group were more child focused and responsive than mothers in the control group.12 One possible explanation for the beneficial effects of the home intervention on academic performance and schoolwork habits at age 8 is that mothers continued to provide responsive and stimulating home environments for their children. Our findings are consistent with other long-term studies that have reported beneficial effects of early home visiting on children’s intellectual performance, vocabulary, arithmetic performance, and behavior.30,46
The significant linear trend in both height and BMI, with children in the FTT-HI group occupying an intermediate position between the children in the FTT-CO and AG groups, suggests that the home intervention may have provided some protection against the relatively poor growth of the children in the FTT-CO group. The mechanism linking home intervention and growth is not clear because the intervention provided neither supplementation nor counseling regarding nutrition or feeding. One possibility is that the sensitivity demonstrated by mothers in the home intervention group may have enabled them to provide a more responsive and interactive feeding environment than mothers in the clinic only group.

When academic skills were considered, children in the AG group outperformed children in the FTT-CO group by an average of ~6.7 points in arithmetic and 4.2 points in reading. The difference in arithmetic was significant, but the difference in reading was not. These findings are alarming because they demonstrate the vulnerability of children with early growth failure that occurs in the midst of urban poverty. The intermediate position of children in the FTT-HI group provides additional evidence that home intervention reduced the negative effects on arithmetic experienced by children with FTT who did not receive home intervention.

The teachers who rated the children’s behavior viewed the children in the FTT-HI group as demonstrating better work habits and fewer behavior problems than children in the FTT-CO group. The teachers were unaware of the children’s growth or intervention history. One possible explanation for this finding is that the intervention group children were exposed to a stimulating home environment that addressed their emotional needs and helped them learn to regulate their behavior and take advantage of learning opportunities at school. These findings are consistent with long-term benefits of early intervention reported from the Abecedarian Project in North Carolina, a home-visiting project among stunted preschool children in Jamaica, the Infant Health and Development Project among low birth weight and premature infants, and a nurse home-visiting program in New York.

Methodologic Considerations
Although this follow-up study avoided many of the sampling and methodologic problems present in follow-up studies of children with FTT that rely on referred or referred children, there are several methodologic considerations. First, the sample is limited to children recruited from pediatric primary care clinics serving a low-income, largely minority community and, therefore, findings are limited to low-income, urban, minority children.

Second, as with most studies, we have suggestive, but not conclusive, evidence regarding the mechanisms linking caregiver behavior and children’s growth, cognitive/academic performance, and behavior. We presume that parenting characterized by responsiveness and stimulating opportunities enabled children to take advantage of developmentally stimulating and educational opportunities.

Clinical Implications
The findings from this investigation generate several recommendations. First, FTT continues to serve as an early warning sign of children’s vulnerability. Although many children with FTT experience growth and cognitive recovery by school age, they continue to be at risk for poor growth, low academic achievement, and poor academic work habits that are likely to undermine future performance. Community studies indicate that up to 50% of children with FTT are not identified, suggesting that the prevalence of FTT is higher than reported estimates. When FTT co-occurs with other threats to children’s well-being, such as neglect, the negative effects on children’s behavior, cognitive, and academic function are compounded. Growth screening strategies to identify children with FTT and interdisciplinary interventions have been effective in promoting children’s growth and development and should be continued. Second, there are few longitudinal studies of children with a history of FTT and most are compromised by serious methodologic problems, including referral bias. Following children with a history of FTT through elementary school and into adolescence would enable investigators to examine how children handle the increasingly sophisticated academic and social challenges that occur during later childhood and adolescence.

Third, the negative consequences of poverty were apparent across multiple indicators of the children’s development, regardless of their growth and intervention status. In communities where there are social safeguards through public assistance programs, such as WIC (Supplemental Nutrition Program for Women, Infants, and Children) and the Food Stamp program, the likelihood of severe FTT is reduced. Strategies are needed to protect children from the negative consequences of poverty through economic resources and opportunities for their families to provide responsive and stimulating caregiving environments.

Finally, early home intervention mitigated many of the negative effects of FTT. It is likely that the stimulating caregiving environment resulting from early intervention was effective in helping children build strong work habits that enabled them to take advantage of academic opportunities. As Nobel laureate and economist James J. Heckman concluded in a recent review, “early interventions targeted toward disadvantaged young children have much higher returns than later interventions.” Efforts to provide early interven-
tion to vulnerable children and their families should be continued, along with long-term follow-up evaluations to assess and ameliorate additional developmental risks.

ACKNOWLEDGMENTS
This research was supported by grants MCJ-240568 and MCJ-240621 from the Maternal and Child Health Research Program, US Department of Health and Human Services, and grants to the Consortium for Longitudinal Studies on Child Abuse and Neglect (LONGSCAN) from the Children’s Bureau, Office on Child Abuse and Neglect, Administration for Children, Youth, and Families.

REFERENCES
35. Bradley RH, Caldwell BM. The HOME Inventory and family demographics. Dev Psychol. 1984;20:315–320


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.
Preterm Infants as Young Adults: A Swedish National Cohort Study

Karolina Lindström, MD, Birger Winbladh, MD, PhD, Bengt Haglund, PhD, Anders Hjern, MD, PhD

OBJECTIVE. Increasing numbers of infants born preterm survive into adulthood. In this study, we analyzed the effect of having been born preterm on disability and vocational success in young adults.

METHODS. A Swedish national cohort of 522,310 infants born in 1973–1979 were followed up for disabilities and income in national registers in 2002 at the age of 23 to 29. Hypotheses were tested in multivariate analysis with logistic regression models on the log scale for dichotomized outcomes and linear regression for continuous variables.

RESULTS. There was a stepwise increase in disability in young adulthood with increasing degree of preterm birth. A total of 13.2% of children born at 24 to 28 weeks’ gestation and 5.6% born at 29 to 32 weeks’ gestation received economic assistance from society because of handicap or persistent illness compared with those born at term after adjustment for socioeconomic and perinatal confounders. Moderate (33–36 weeks’ gestation) and marginal (37–38 weeks’ gestation) preterm birth also carried significantly increased risks for disability and were responsible for 74% of the total disability associated with preterm birth. Preterm birth was associated with a lower chance of completing a university education and a lower net salary in a stepwise manner. The total economic gain for Swedish society, in terms of taxes and decreased costs for benefits, if all long-term effects of preterm birth could have been prevented in the birth cohorts in this study, would have amounted to 65 million euros in 2002 alone.

CONCLUSIONS. The majority of adults who were born very preterm lived an independent and self-supportive life. Moderately preterm birth carries a considerable risk for long-term impairment. There are strong economic incentives for secondary prevention of disability associated with preterm birth.
Increasing numbers of children born preterm survive into adulthood as a consequence of progress in prenatal and neonatal care. A considerable number of cohorts of former patients from NICUs have been created and followed prospectively over time. At school age 10% to 12% have been described to have considerable impairment because of a neurologic disability, as well as low cognitive test scores and an increased risk of behavior disturbances including attention-deficit/hyperactivity disorder. A recent follow-up at 20 years of age demonstrated a slightly lower IQ at 20 years of age and a persistent educational disadvantage in preterm with a birth weight <1500 g. Impaired cognitive function and lower academic performance at school age seem to persist into young adulthood, although a majority seem to have overcome earlier difficulties at the age of 22 to 25 years.

Despite the abundance of data regarding test results and disorders of extremely preterm children, the overall importance of these minor and major disabilities for their educational and vocational career as adults is yet to be described. Also, few follow-up studies have addressed the situation of the much larger group of children born moderately preterm. In this study, we used the Swedish national databases to study social outcomes in an entire national cohort of young adults, 23 to 29 years of age, in relation to gestational age.

METHODS
This study was based on Swedish national registers held by the National Board of Health and Welfare and Statistics Sweden linked through each individual’s unique personal identification number. A total of 570,768 live-born individuals were identified in the Swedish medical birth register 1973–1979 with a reported gestational age ≤41 weeks. In this register age, we also identified the personal identification number of the mother, geographic location of the household, birth date and gender of the child, and major perinatal indicators (gestational age, Apgar <7 at 5 minutes, neonatal distress, birth weight, small for gestational age [SGA], preeclampsia, abruptio placentae, preterm rupture of the membranes). Gestational age was defined according to maternal report of last menstrual period and clinical judgment by the attending pediatrician. SGA was defined as less than −2 SD according to the growth chart developed by Marsal et al. A total of 4022 (0.7%) children with a registered birth weight >3 SD or less than −6 SD according to this same growth chart were excluded from the study population as probable coding errors, because these errors are more probable with birth weights categorized as large for gestational age than in children categorized as SGA in preterm infants. A total of 12,927 (2.3%) individuals with at least 1 reported malformation other than undescended testicle, preauricular appendage, congenital nevus, or hip dislocation were also excluded, as were 7310 (1.3%) individuals who were reported dead in the Swedish national cause-of-death register by 2002. The infant (1 year) mortality according to gestational age was 43.4% in weeks 24 to 28, 13.6% in weeks 29 to 32, 1.6% in weeks 33 to 36, 0.38% in weeks 37 to 38, and 0.17% in weeks 39 to 41 when infants with major malformations were excluded. Finally, we excluded 27,041 persons (4.7%) who were no longer residents in Sweden in 2002 according to the register of the total population in December 2002, leaving 522,310 individuals with a gestational age between 24 and 41 weeks to be included in the study population.

The socioeconomic status (SES) of the household, housing situation, maternal country of birth, and lone parent households were identified in the Swedish Population and Housing Census of 1985. Socioeconomic groups were defined according to a classification created by Statistics Sweden (manual workers, skilled workers, and white collar 1–3). This hierarchical classification of occupation is based on the educational level required but also takes type of work and the position at the work place into consideration. Farmers, self-employed and unemployed, were categorized as unclassified. The head of the household was defined as the adult in the household with the highest SES. Maternal country of birth was categorized into 4 geographical groups: Sweden, Finland, the rest of Europe, and the rest of the world. Social welfare benefits received by the household of the mother were added through linkage to the Total Enumeration Income Survey for 1990.

Social Outcomes
Several social outcome variables were created with 2002 information from the Total Enumeration Income Survey that year: (1) sickness pension, indicating lifelong pension because of long-standing illness or disability, (2) handicap allowance, indicating a permanent disability, (3) employment, indicating having an income from employment or own firm in November 2002, (4) illness benefits, indicating temporary economic support during at least 2 consecutive weeks because of illness from the national health insurance, (5) residence in the household of the biological parents, (6) student, as indicated by having received student support or loan, (7) net salary, indicating net income from employment and own firm, (8) disposable income, indicating the sum of all incomes, including societal benefits, deducted by income tax, and (9) net transfer to society, as indicated by deducting disposable income from work income. Economic compensation because of disability assistance, indicating the need for a personal helper at least 4 hours daily, was identified in the register kept at the National Social Insurance Board.

The number of months that social assistance was received during 2002 was identified in the Swedish social-assistance register, and the highest completed education
as of December 2002 was derived from the Swedish educational register. Education was categorized as “basic” if the study participant had completed no more than the compulsory 9 years of primary school, and as “post-secondary” if ≥1 educational level had been completed after secondary school. A summarized disability outcome variable was created that indicated having received sickness pension, handicap allowance, and/or disability assistance.

Statistical Methods
Multivariate analyses of the dichotomised social outcome variables described above were calculated by logistic regression on the log scale to calculate estimates equivalent to relative risk ratios (RRs). We calculated 95% confidence intervals by using the test-based method. Disability and residence in the household of the biological parents were analyzed in the entire study population. Disabled individuals, however, were excluded from the analysis of educational outcomes to allow for an analysis of more subtle consequences. Employment, finally, was studied in those in the study population that had neither an indication of disability nor of being a student.

Year of birth was entered as a continuous variable in the regression models. Missing data were entered as a separate category in the analytic models. Model 1 in the logistic regression analysis was adjusted for gender and age only. Socioeconomic and parental morbidity variables were added as confounders in model 2. SGA and multiple birth were added to the variables of model 2 as perinatal confounders in model 3. Interaction effects were tested in logistic regression models. Attributable risk was calculated according to the formula suggested by Rothman and Greenland.

Income variables were analyzed in a linear regression with the same independent variables as in the logistic regression models described above. Students were excluded from this analysis, and only those employed without a disability were included in the analysis of net salary. All statistical analyses were conducted by using SAS 9.0 software (SAS Institute, Inc, Cary, NC).

RESULTS
There were 431 656 individuals who were born term (39–41 weeks’ gestational age), 68 541 born slightly preterm (37–38 weeks’ gestational age), 19 166 born moderately preterm (33–36 weeks’ gestational age), 2947 born very preterm (317 born 24–28 weeks’ gestational age, and 2630 born 29–32 weeks’ gestational age) in the study population. Among the very preterm infants, 43 were born with a gestational age of ≤26 weeks.

Sociodemographic data of the study population by gestational age are presented in Table 1. Boys were more often born very (24–32 weeks’ gestational age) and moderately preterm (33–36 weeks’ gestational age). In 1985, the households of the study subjects born preterm more often lived in rented apartments, received social welfare, and had an SES as manual workers compared with the households of those born at term.

Perinatal variables by gestational age are presented in Table 2. Perinatal and gestational complications increased by increasing degree of preterm birth.

The rate of disability in the entire study population decreased by year of birth from 1.66% in those born 1973 to 1.46% in those born 1979; a RR of 0.97 (95% confidence interval [CI]: 0.96–0.98) per year. Social outcomes in relation to gestational age at birth are presented in Table 3. Rates of disability allowance, sickness pension, and disability assistant increased gradually in step with the degree of preterm birth. More than 85% of even the extremely preterm children (24–28 weeks’ gestation), however, had no such indication of a disability. After adjustment for age and gender in model 1 in the multivariate analysis, the RRs of children born very preterm was 4.39 (Table 4) for having a disability. Adding socioeconomic and perinatal confounders decreased the RRs slightly to 3.76. Moderately preterm (33–36 weeks’ gestational age) and slightly preterm (37–38 weeks’ gestational age) had RRs of 1.51 and 1.23, respectively, in the final model. The total attributable risk for disability associated with preterm births was 7.0%, of which 5.2% was accounted for by moderately and slightly preterm births. Having received social welfare benefits and sick allowance had a distribution by gestational age that was similar to the disability indicators.

Study subjects born preterm more often resided in the household of their parents compared with peers of the same gender and age who were born term (RR: 1.15 [95% CI: 1.05–1.25], 1.05 [95% CI: 1.01–1.09], and 1.03 [95% CI: 1.01–1.05]) for extremely, moderately, and slightly preterm groups, respectively (data not shown in the tables). Adding confounders only marginally changed these estimates.

A total of 26% of those born with a gestational age of 28 weeks or less had a university education in comparison with 38% of those with a normal gestational age (39–41 weeks). When disabled individuals were excluded, this difference decreased to an adjusted RR of 0.93 (Table 4) for having a university education. Moderately preterm individuals had a RR of 0.96 in the final model, whereas slightly preterm individuals had a similar chance as those born term to have a university education in the final model (Table 4).

A total of 74.1% of the study population born at term had an income from employment in November 2002 compared with 68.1% of those born in gestational weeks 24 to 28, 70.1% in weeks 29 to 32, and 72.5% of those born in weeks 33 to 36. When disabled individuals were excluded, however, preterm and term individuals had very similar chances of having been employed (Table 4).
As Table 5 demonstrates, very preterm birth was associated with a lower net salary among those who were employed in a linear regression analysis of income, even after disabled individuals were excluded and socioeconomic and perinatal confounders were accounted for. Adults born preterm also had a lower disposable income than adults born term (Table 5). When the disposable income was subtracted from the net salary, a net mean individual transfer of 3079 euros to the society from the individuals born term in the study population was identified in 2002 (Table 5). In a linear regression, it was demonstrated that this transfer was 715 euros lower \((P = .02)\) in those born very preterm and 171 euros lower \((P = .15)\) in those born moderately preterm after socioeconomic and perinatal confounders were accounted for (Table 5). In the entire population, this amounts to a net negative transfer of ~65 million euros from individuals born preterm compared with those born term in 2002 alone, with those born moderately preterm accounting for 77\% of that amount.

Differences in effects of preterm birth of gender and SES on university education, employment, and disability were studied in interaction analyses with the variables in model 1 and a dichotomized SES variable in the logistic regression analysis presented above. The effect of preterm birth on the chance of receiving a university education was greater in families with a low SES compared with a high SES \((P = .005)\); otherwise, no statistically significant interaction effects for SES or gender were found.

**DISCUSSION**

In this national cohort study of Swedish residents born in the 1970s, we have demonstrated that preterm birth is associated with educational and vocational impairment in young adulthood. As expected, the effect was greatest for the 0.6\% born very preterm (24–32 weeks' gestational age), but even in this group, the large majority were employed and contributed more in income tax than they received in benefits. Somewhat more unexpected were the considerable effects observed in the 3.7\% who were born moderately preterm.

Preterm birth was associated with a stepwise increased risk of disability but also with a lower net income.
and a lower chance of completing a university education than in those without any indication of a disability. A number of obstetrical and neonatal complications, such as intrauterine growth retardation, cerebral hemorrhage and infarction, hypoglycemia, pulmonary disease, septicemia, and asphyxia, may cause long-term disability through injuries to the immature central nervous system and are more common in preterm births. It seems less certain that these catastrophic injuries are responsible for the more subtle educational and vocational impairment of preterm children without an identified disability. It is possible that mechanisms such as incomplete myelinization and poor growth of the gray matter of the preterm brain have to be considered to explain these effects.

This study demonstrates an attributable risk of 7% for...
preterm birth on disability in this cohort of young adults, in which moderately and slightly preterm births accounted for 5%. In a similar manner, moderately preterm accounted for 61% of the societal transfers to individuals born preterm. These findings are in accord with recent studies that have described considerable mortality and respiratory morbidity in infants born at 30 to 34 weeks’ gestation,8 as well as in those with 35 to 36 weeks’ gestation.8 The results of this study support calls for more research and development of improved secondary preventive strategies in this group of infants.

The analysis of income and societal transfers in this study indicates that the Swedish society as a whole has gained considerably from the increased survival and productivity in vulnerable neonates born preterm associated with the development of neonatal intensive care. Even very preterm infants made such economic contributions to the society in young adulthood (mean: 2000 euros/year) that it seems most probable that this would hold true even if other societal expenses, such as increased needs for medical care and drugs, were added to the bill. Compared with children born term, however, children born preterm contributed less to society. This demonstrates that even quite ambitious and costly primary and secondary preventive strategies are needed for this group of infants.

### Table 4: Logistic Regression of Social Outcomes

<table>
<thead>
<tr>
<th>Outcome</th>
<th>N</th>
<th>Model 1, RR (95% CI)</th>
<th>Model 2, RR (95% CI)</th>
<th>Model 3, RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disabilitya</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24–32 wk gestation</td>
<td>2947</td>
<td>4.39 (3.79–5.07)</td>
<td>4.11 (3.44–4.90)</td>
<td>3.76 (3.14–4.49)</td>
</tr>
<tr>
<td>33–36 wk gestation</td>
<td>1916</td>
<td>1.86 (1.70–2.03)</td>
<td>1.76 (1.58–1.97)</td>
<td>1.51 (1.32–1.73)</td>
</tr>
<tr>
<td>37–38 wk gestation</td>
<td>6854</td>
<td>1.28 (1.20–1.36)</td>
<td>1.26 (1.18–1.36)</td>
<td>1.26 (1.17–1.35)</td>
</tr>
<tr>
<td>39–41 wk gestation</td>
<td>431656</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Postsecondary educationb</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24–32 wk gestation</td>
<td>2758</td>
<td>0.87 (0.82–0.93)</td>
<td>0.92 (0.85–0.98)</td>
<td>0.92 (0.86–0.99)</td>
</tr>
<tr>
<td>33–36 wk gestation</td>
<td>18645</td>
<td>0.91 (0.89–0.94)</td>
<td>0.97 (0.94–0.99)</td>
<td>0.96 (0.93–0.99)</td>
</tr>
<tr>
<td>37–38 wk gestation</td>
<td>67266</td>
<td>0.96 (0.97–0.99)</td>
<td>0.99 (0.98–1.01)</td>
<td>1.00 (0.98–1.01)</td>
</tr>
<tr>
<td>39–41 wk gestation</td>
<td>425288</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Model 1 is adjusted for age and gender; model 2 is adjusted for age, gender, SES, and single-parent household in 1985 and residency, maternal age, parity, and social welfare in 1990; and model 3 is adjusted for age, gender, SES, and single-parent household in 1985 and residency, maternal age, parity, and social welfare in 1990 and parental psychiatric disorders, SGA, and multiple birth.

a Analysis includes entire study population.
b Analysis excludes individuals with at least 1 indication of a disability.

### Table 5: Linear Regression of Income Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean, Euro</th>
<th>Median, Euro</th>
<th>Model 1 B</th>
<th>P</th>
<th>Model 2 B</th>
<th>P</th>
<th>Model 3 B</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Net salarya</td>
<td></td>
<td></td>
<td></td>
<td>-732</td>
<td>.000</td>
<td>-488</td>
<td>.041</td>
<td>-503</td>
<td>.036</td>
</tr>
<tr>
<td>24–32 wk gestation</td>
<td>1749</td>
<td>22 271</td>
<td>22 894</td>
<td>-1071</td>
<td>.000</td>
<td>-876</td>
<td>.003</td>
<td>-825</td>
<td>.005</td>
</tr>
<tr>
<td>33–36 wk gestation</td>
<td>12043</td>
<td>22 735</td>
<td>23 370</td>
<td>-326</td>
<td>.000</td>
<td>-332</td>
<td>.005</td>
<td>-312</td>
<td>.009</td>
</tr>
<tr>
<td>37–38 wk gestation</td>
<td>42051</td>
<td>22 861</td>
<td>23 448</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Disposable incomeb</td>
<td></td>
<td></td>
<td></td>
<td>-910</td>
<td>.002</td>
<td>-685</td>
<td>.021</td>
<td>-715</td>
<td>.016</td>
</tr>
<tr>
<td>24–32 wk gestation</td>
<td>2238</td>
<td>15 749</td>
<td>15 859</td>
<td>-1015</td>
<td>.000</td>
<td>-987</td>
<td>.003</td>
<td>-965</td>
<td>.003</td>
</tr>
<tr>
<td>33–36 wk gestation</td>
<td>14521</td>
<td>16 365</td>
<td>16 519</td>
<td>-489</td>
<td>.000</td>
<td>-332</td>
<td>.005</td>
<td>-312</td>
<td>.009</td>
</tr>
<tr>
<td>37–38 wk gestation</td>
<td>50512</td>
<td>16 599</td>
<td>16 663</td>
<td>-236</td>
<td>.000</td>
<td>-155</td>
<td>.019</td>
<td>-152</td>
<td>.021</td>
</tr>
<tr>
<td>39–41 wk gestation</td>
<td>320654</td>
<td>16 861</td>
<td>16 840</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Model 1 is adjusted for age and gender; model 2 is adjusted for age, gender, SES, residency, maternal age, parity, and social welfare in 1990; and model 3 is adjusted for age, gender, SES, residency, maternal age, parity, and social welfare in 1990 and parental psychiatric disorders, SGA, and twins/triplets.

a Analysis excludes individuals with at least 1 indication of a disability and students.
b Analysis excludes students.
secondary preventive programs in perinatal health have a potential of economic gains for the society in the long run. The potential economic gain for the Swedish society, in the cohorts in this study if all long-term effects of preterm birth could have been prevented would have been ~65 million euros in 1 year alone in terms of taxes and decreased costs for benefits.

The major strength of this study is the large study population made possible by the high quality and extensive coverage of the Swedish national registers, which enabled us to study a large number of infants with marginal attrition. The use of register data on income, disability, and education excludes the possibility of self-reporting bias, which is otherwise a common problem in follow-up studies. We were able to include a wide range of socioeconomic, perinatal, and parental morbidity variables as possible confounding factors in the analysis by linking several national registers. It is still quite possible, however, that important factors associated with preterm birth that might also affect the fetus directly, such as viral infections or toxic agents, were not accounted for in the multivariate analysis.

The cohort in this study was born before ultrasound became a routine procedure to measure gestational age in early pregnancy. Although a certain socioeconomic bias in coding may be inherent with the variable of gestational age used in this study, most of this bias can be accounted for in the multivariate analysis. Coding errors of gestational age tend to create outliers by linking several national registers. It is still quite probable, however, that important factors associated with preterm birth that might also affect the fetus directly, such as viral infections or toxic agents, were not accounted for in the multivariate analysis.

The major strength of this study is the large study population made possible by the high quality and extensive coverage of the Swedish national registers, which enabled us to study a large number of infants with marginal attrition. The use of register data on income, disability, and education excludes the possibility of self-reporting bias, which is otherwise a common problem in follow-up studies. We were able to include a wide range of socioeconomic, perinatal, and parental morbidity variables as possible confounding factors in the analysis by linking several national registers. It is still quite possible, however, that important factors associated with preterm birth that might also affect the fetus directly, such as viral infections or toxic agents, were not accounted for in the multivariate analysis.

A recent analysis of cerebral palsy in neonates with a very low birth weight in 16 different European cities demonstrates decreasing rates in 1996 compared with 1980 despite higher survival rates of infants with very low birth weight in 1996. The lower rate of cerebral palsy was particularly noted among children born in gestational weeks 29 to 32, whereas the rate remained quite similar over time in children born at 24 to 28 weeks’ gestation. Swedish studies have demonstrated a similar pattern of improvement. Thus, we may expect that disability rates of preterm with a gestation of 29 to 32 weeks born in Sweden during more recent years will have even more favorable outcomes than those reported in this study of cohorts from the 1970s.

CONCLUSIONS
This study demonstrates that there is considerable long-term impairment associated with preterm birth. This impairment generates great costs for the society, although the large majority of even the most immature children seemed to live a productive and self-supported life in early adulthood. More attention is needed toward the moderately preterm children who accounted for most of the chronic disability associated with preterm birth in this population of young adults.

ACKNOWLEDGMENT
This work was supported by the National Board of Health and Welfare, Stockholm, Sweden.

REFERENCES
7. Escobar GJ, Clark RH, Greene JD. Short-term outcomes of infants born at 35 and 36 weeks gestation: we need to ask more questions. Semin Perinatol. 2006;30:28–33

IN PRAISE OF LESS PRAISE

“For management consultant Jerry Pounds, toasters were a turning point. He had built a lucrative career advising companies on ways to praise employees, especially younger ones, who grew up bombarded with soccer trophies, parental applause and stroking at school. Mr Pounds figures he trained 50,000 supervisors, encouraging them to pass out cartloads of prizes, plaques and praise-engraved knickknacks. Then several nurses who received toasters as incentives told him they were insulted. ‘I got into nursing to care for patients,’ one said, ‘not so I’d be rewarded with toasters.’ Mr Pounds says he came to some realizations: Unearned praise is condescending and destructive, incentives become entitlements and ‘we’ve ruined our kids’ by celebrating mediocrity. Mr Pounds contacted me in response to my recent Weekend Journal article on praise in today’s workplace. It explained how companies are celebrating young employees by throwing confetti at them, passing out ‘applause notes’ and giving them kudos just for coming to work on time. The article drew attention from bloggers, talk-radio hosts, and a slew of e-mailing readers. Advice was sharp: Rain on their parades. Many argued that to counterbalance our praise culture, young people need reality slaps. David Dumpe, a professor at Kent State University, now begins each semester by asking students: ‘How many of your parents raised you by saying you can be anything you want to be?’ Two-thirds raise their hands, he says. He then asks: ‘Do you realize that’s a bunch of baloney?’ In Iowa, a teacher tells incoming seventh graders: ‘Your entire life you have heard from parents that you are wonderful—the center of the universe. It’s not true. You are not wonderful. You are one of many.’ In part because of his refreshing bluntness, this teacher is beloved by students, a colleague writes.”


Noted by JFL, MD
Three-Tesla Cardiac Magnetic Resonance Imaging for Preterm Infants

Adrienne M. Foran, MD\textsuperscript{a,b}, Julie A. Fitzpatrick, BS\textsuperscript{b,c}, Joanna Allsop, BS\textsuperscript{b,c}, Stephan Schmitz, PhD\textsuperscript{b,c}, Jamie Franklin, BS\textsuperscript{b,c}, Constandinos Pamboucas, MD\textsuperscript{b,c}, Declan O’Regan, FRCR\textsuperscript{b,c}, Jo V. Hajnal, PhD\textsuperscript{b,c}, A. David Edwards, FMedSci\textsuperscript{a,b,c}

Departments of \textsuperscript{a}Paediatrics and \textsuperscript{b}Imaging Sciences, Imperial College, London, England; \textsuperscript{c}Medical Research Council Clinical Sciences Centre, Hammersmith Hospital, London, England

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

ABSTRACT

OBJECTIVES. We aimed to establish the feasibility of acquiring 3.0-T cardiac MRIs without sedation, anesthesia, or breath-holding for preterm infants and to obtain preliminary quantitative data on left ventricular function in this population.

METHODS. Twelve preterm infants underwent 3.0-T cardiac MRI without sedation or breath-holding. The median gestational age was 29 weeks (range: 26–33 weeks), the median birth weight was 1240 g (range: 808–2200 g), and the median postconceptional age at the time of cardiac MRI was 33 weeks (range: 31–40 weeks). Anatomic images were acquired with T2-weighted spin-echo sequences, and ventricular function was assessed with balanced steady-state free precession cine sequences. We assessed left ventricular function by using the area-length ejection fraction method on horizontal long-axis images and the volumetric Sergeant’s discs method of analysis on short-axis images.

RESULTS. Imaging was successful for 10 of 12 infants. For those 10, the area-length ejection fraction method in the horizontal long-axis plane estimated median stroke volume at 2.9 mL, cardiac output at 0.4 L/minute, end-diastolic volume at 3.8 mL, end-systolic volume at 0.3 mL, and ejection fraction at 74.6%. Short-axis volumetric estimations were made for 4 infants. With this approach, the median stroke volume was 2.4 mL, cardiac output 0.35 L/minute, end-diastolic volume 4.3 mL, end-systolic volume 2.1 mL, and ejection fraction 56%.

CONCLUSIONS. Three-tesla cardiac MRI is feasible for preterm infants without sedation, anesthesia, or breath-holding and has the potential to provide a wide range of precise quantitative data that may be of great value for the investigation of cardiac function in preterm infants.
CARDIAC FUNCTION in preterm infants is poorly understood. Commonly used bedside tests such as blood pressure and capillary refill time have very low specificity and sensitivity in detecting low blood flow in preterm infants. There is emerging evidence that cardiovascular function is an important determinant of outcome. Low superior vena cava flow is common in the first hours after birth and has been associated with subsequent periventricular hemorrhage, whereas inflammatory mediators induced by sepsis impair cardiac function and low cardiac output is associated with reduced cerebral blood flow.

Most available data on cardiac function in preterm infants have been acquired by using echocardiography, which is an invaluable tool for observing cardiac morphologic features but is relatively imprecise as a measure of function in individuals. Echocardiographic approaches have been unable to provide definitive answers to important practical questions, such as how cardiac output is related to blood pressure and how volume expansion and inotropic therapy affect cardiac output in extremely preterm infants.

Cardiac MRI (CMRI) has been a major advance contributing to the understanding of cardiac function and disease in adults, providing quantitative data with high levels of accuracy and reproducibility. However, CMRI is technically challenging in children because of small cardiac size, rapid heart rates, and a high level of sensitivity of CMRI to motion-induced degradation of image quality. Furthermore, in previous CMRI studies, children were usually sedated or anesthetized and established on mechanical ventilation to overcome movement and to allow breath-holding, a strategy that yields significant improvements in image quality in adults. Unfortunately, sedation or anesthesia may have particular problems in preterm infants.

To our knowledge, there have been no previous reports of CMRI in preterm infants undergoing intensive care. The previous studies of CMRI in older infants and children have been at 1.5-T field strength, mirroring the predominance of 1.5-T CMRI for adults and the relative difficulty of developing 3.0-T systems for adult CMRI. However, 3.0 T is becoming widely established as the standard field strength for imaging other parts of the body. Higher field strengths offer the potential advantage of higher signal/noise ratios, allowing improved spatial or temporal resolution once appropriate imaging parameters have been defined. Interestingly, although the small physical size of the preterm heart has been an impediment to development of CMRI in general, it offers a specific advantage in the physical context of higher-field imaging; this suggested to us that development of 3.0-T CMRI might be appropriate for this population.

If appropriate techniques can be developed, CMRI has the potential to provide valuable information on cardiac function in preterm infants and to address long-standing practical problems in neonatal intensive care. The purpose of this study was thus to establish the feasibility of CMRI at 3.0 T in nonsedated preterm infants. We aimed to define imaging parameters and to obtain preliminary quantitative data on left ventricular function by using 2 standard techniques, that is, the area-length ejection fraction (ALEF) approach on the horizontal long-axis view, which is rapid and simple but requires an assumption of ellipsoid ventricular geometry, and the volumetric Sergeant’s discs method, which avoids the ellipsoid assumption but requires a contiguous stack of short-axis images.

METHODS

Subjects

Twelve preterm infants underwent CMRI. Clinical details are given in Table 1. The median gestational age was 29 weeks (range: 26–33 weeks), the median birth weight was 1240 g (range: 808–2200 g), and the median postconceptional age at the time of MRI was 33 weeks (range: 31–40 weeks). For 2 infants, patent ductus arteriosus (PDA) had been diagnosed previously through echocardiography; those infants required additional inspired oxygen. The remaining 10 subjects were thought...
to have structurally and functionally normal hearts and were not oxygen dependent.

Infants were fed and allowed to fall into natural sleep in a quiet environment and then were laid in a custom-made, MRI-compatible cradle and placed in the scanner, with MRI-compatible physiologic monitoring and ear protection, as reported previously. Infants were monitored with electrocardiography, pulse oximetry, and a video camera, and a pediatrician was present throughout the procedure. Each scan required ~45 minutes.

The Hammersmith Hospital research ethics committee granted ethical permission for 3.0-T imaging studies. Written informed parental consent was obtained for each infant studied.

**Image Acquisition**

All scans were conducted with a Philips 3-T Intera system (Philips, Best, Netherlands). A surface receiver coil was placed over the chest wall. Vector electrocardiographic gating was used conventionally, but no motion-correcting pulse sequences or respiratory triggering was used. An axial, electrocardiographically triggered, T2-weighted, black-blood, anatomic sequence was used to give an overview of the heart and great vessels. Automatic shimming was used. Balanced steady-state free precession (b-SSFP) cine sequences were acquired in the horizontal and vertical long-axis planes. These were used to plan a stack of contiguous slices in the left ventricular short-axis plane.

After investigating image acquisition by using varying parameters, we found that, for spin-echo anatomic images, a repetition time of 2 beats and an echo time of 60 milliseconds yielded the highest-quality images. For b-SSFP cine sequences, best tissue contrast was achieved with a flip angle of 45°. Specific absorption rate values, as provided by the scanner, were noted for all examinations, and all infants were monitored for signs of thermal stress. MRI parameters for the reported images are given in Table 2.

**Quantitation of Cardiac Function**

The images were analyzed on a cardiac workstation (Philips View Forum). The first image in each cine sequence was considered to be at end-diastole. The end-systolic phase was determined as the slice with the smallest left ventricular cavity. Endocardial borders were traced manually; papillary muscle was demarcated separately and excluded from the ventricular volume.

Left ventricular volume was measured in 2 ways. First, the ALEF approach was used. A single long-axis view was obtained, and a single-plane ellipsoid model was applied to estimate the ventricular volume, as demonstrated in Fig 1. This technique assumes that the left ventricle has an ellipsoid shape. The left ventricular contour is outlined (green) at end-diastole and end-systole. The ALEF was calculated by using the standard mathematical calculation for volume of an ellipsoid shape, as \( \frac{4}{3}\pi \times \frac{A^2}{L} \), where \( A \) is the cavity area and \( L \) (yellow line) is the cavity long-axis measurement.

Second, by using the Sergeant’s disk approach, a series of short-axis views were taken to encompass the left ventricular volume, and the left ventricular volume was calculated as the sum of each short-axis area multiplied by the slice thickness. Slices were considered to be within the left ventricle if \( \geq 50\% \) of the ventricular myocardium surrounded the blood volume.

The following left ventricular parameters were then estimated, by using both techniques, for infants for whom appropriate images could be obtained: ventricular ejection fraction (milliliters), ventricular end-diastolic and end-systolic volumes (milliliters), ventricular stroke volume (milliliters), and cardiac output (liters per minute). Test-retest consistency was assessed by calcu-

---

**TABLE 2**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Balanced Field Echo (b-SSCP)</th>
<th>T2-Weighted Black Blood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flip angle, degrees</td>
<td>45</td>
<td>90</td>
</tr>
<tr>
<td>Repetition time, ms</td>
<td>3–4</td>
<td>2</td>
</tr>
<tr>
<td>Echo time, ms</td>
<td>1.5–2</td>
<td>60</td>
</tr>
<tr>
<td>Matrix, pixels</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency</td>
<td>256</td>
<td>256</td>
</tr>
<tr>
<td>Phase</td>
<td>144</td>
<td>256</td>
</tr>
<tr>
<td>Field of view, mm</td>
<td>160</td>
<td>180</td>
</tr>
<tr>
<td>Slice thickness, mm</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Echo train length</td>
<td>NA</td>
<td>14</td>
</tr>
<tr>
<td>No. of signal averages</td>
<td>1 or 2</td>
<td>3</td>
</tr>
<tr>
<td>Specific absorption values, W/kg</td>
<td>3.8</td>
<td>3.8</td>
</tr>
</tbody>
</table>

NA indicates not applicable.

---

**FIGURE 1**

Left ventricular, long-axis CMRI scans of the heart in 2 preterm infants, 1 with a structurally normal heart (A and B) and 1 with a large PDA (C and D). The left ventricular contour is outlined (green) at end-diastole (A and C) and end-systole (B and D). We used the standard mathematical formula for calculating ellipsoid volume, that is, \( \frac{4}{3}\pi \times \frac{A^2}{L} \), where \( A \) is the cavity area and \( L \) is the cavity length (yellow line). These long-axis images allow estimations of ventricular volume, stroke volume, and cardiac output.
lating the coefficient of variation for 10 repeated determinations for 1 subject with both long-axis and short-axis methods. The mean difference between the long-axis and short-axis methods was estimated through calculation of the mean difference, according to the approach described by Bland and Altman.18

RESULTS
For 10 of 12 subjects, images were obtained that allowed long-axis (ALEF) estimations; for 4 of those infants, short-axis (Sergeant’s discs) calculations were also performed. For the remaining 2 infants, movement artifacts prevented image analysis. Examples of images at end-systole and end-diastole with these 2 methods of analysis are shown in Figs 1 and 2. A series of images taken from the b-SSFP cine sequence for the cardiac cycle from diastole to systole are given in Fig 3.

Values for ventricular volume, cardiac output, and ejection fraction, as estimated with both methods, are given in Figs 4 and 5. We also calculated the values related to the infants’ weights at the time of scanning. For the ALEF method, the mean values were as follows: stroke volume, 1.6 mL/kg (range: 0.4–3.5 mL/kg); cardiac output, 0.25 L/minute per kg (range: 0.5–0.55 L/minute per kg); end-diastolic volume, 2.2 mL/kg (range: 0.6–4.6 mL/kg); end-systolic volume, 0.6 mL/kg (range: 0.1–1.1 mL/kg). For the Sergeant’s discs method, the mean values were as follows: stroke volume, 1.3 mL/kg (range: 0.8–2.0 mL/kg); cardiac output, 0.18 L/minute per kg (range: 0.13–0.29 L/minute per kg); end-diastolic volume, 2.3 mL/kg (range: 1.5–3.5 mL/kg); end-systolic volume, 1.0 mL/kg (range: 0.8–1.5 mL/kg). The 2 infants with previously diagnosed PDA had much larger left ventricles and showed strikingly increased left ventricular output, achieved primarily by an increase in end-diastolic volume, with similar ejection fractions.

The coefficients of variation for repeated estimations of end-diastolic volume by using long-axis and short-axis methods were 6% and 3%, respectively, and those for end-systolic volume were 7% and 5%, respectively. The 2 methods could be compared for 4 infants (8 measurements). In all except 1 of those comparisons, the short-axis method yielded a slightly higher value, compared with the long-axis approach. The mean ± SD difference (short axis – long axis) for end-diastolic volume was 1.075 ± 0.2 mL, and that for end-systolic volume was 0.3 ± 0.1 mL. Higher end-diastolic volume led to systematically lower estimates for stroke volume and ejection fraction with the short-axis approach (Figs 4 and 5). The indicated specific absorption rate values (3–4 W/kg) were within safety guidelines for all examinations, and none of the infants showed any signs of thermal stress.

DISCUSSION
This study demonstrates that high-quality CMRI at 3.0-T field strength can be performed for preterm infants without sedation, anesthesia, or breath-holding. In the adult population, the excellent cardiac observation offered by CMRI allows a marked improvement in reproducibility for parameters of ventricular function, compared with echocardiography,9 which translates into both more-accurate clinical decision-making and reduced sample sizes for modeling, hemodynamic, and experimental studies.16 Most CMRI experience to date has been gained at 1.5 T, although there is increasing interest in the use of 3.0-T systems, which are becoming widely used for other body MRI applications.17 Imaging at 3.0 T can provide a potential doubling of the signal/noise ratio, compared with conventional 1.5-T platforms.13,19 At present, 3-T CMRI for adults remains challenging, largely because of the technical difficulties inherent with the higher field strength. Although the signal/noise ratio is higher, there are increased problems of signal stability during respiration, difficulty in shimming the static (B0) magnetic field, and inhomogeneity of the radiofrequency (B1) magnetic fields. Susceptibility effects increase linearly with B0 and so double between 1.5 T and 3 T. The air spaces in the lungs pose greater problems at higher field, but this may be less of an issue with neonates, who tend to have higher lung water content and smaller air spaces. The relatively small wavelength, compared with the dimensions of the adult thorax, is associated with B1 inhomogeneity, which is substantially more pronounced at 3 T than at 1.5 T.20 However, the smaller size of neonates greatly reduces this problem. The smaller size of the neonatal heart also puts a premium on the signal/noise ratio, an advantage of 3 T over 1.5 T. Therefore, neonates are

FIGURE 2
Series of short-axis CMRI scans through the ventricle from apex to base. With images obtained at end-diastole (A–F), delineation of the endocardial margin of the ventricle (green) allows the surface area of each slice of ventricle to be calculated and values summed over the set of slices, for estimation of ventricular volume. Comparison with similar measurements made at end-systole allows calculation of ventricular output (G–L).
likely to be a favorable subject group for 3.0-T imaging. The fact that high-temporal resolution, cardiac, cine imaging can be obtained at 3.0 T is important for accurate measurement of functional parameters.9,13

Because CMRI is highly sensitive to movement artifact, breath-holding protocols are used commonly in adult studies to obtain high-quality images. In pediatric studies, breath-holding is often enforced and movement is prevented through the induction of general anesthesia, with mechanical ventilation. Because many neonatologists now aim to avoid mechanical ventilation for preterm patients whenever possible, this approach is unattractive; if breath-holding were required for successful CMRI, it is unlikely that the technique could be used widely in neonatal care or research. We found that a simple regimen of prefeeding in a quiet environment and good acoustic protection allowed us to obtain high-quality images for 10 of 12 subjects, although gross movement prevented image acquisition suitable for analysis for 2 infants. Scanning was usually completed within 1 hour, including the time needed to ensure that the infant was asleep. These findings show that CMRI is a practical technique for the preterm population.

This is the first study to attempt 3.0-T CMRI for preterm infants, and the preliminary data obtained are not intended to define reference ranges for estimated values. At present, that are few data using CMRI to establish references values for the pediatric population, and no published normal values are available for the preterm population.21 This study provides the correct imaging parameters to allow high-quality CMRI for preterm patients at 3.0 T. The number of images presented is appropriate to achieve this goal. With the small number of infants in this study, we did not look for predictable effects, such as the dependence of cardiac output on age, and we did not formally compare these results with echocardiograph estimations; additional studies should be able to acquire the relevant data.

We obtained images of sufficient quality with both long-axis (ALEF) and short-axis (Sergeant’s discs) methods. In this preliminary study, we did not aim for a formal assessment of the comparative values of the ALEF and Sergeant’s discs analyses. It seems that the ellipsoid...
assumption of the ALEF approach may need to be questioned for preterm infants. Calculation of volumes from short-axis images does not require any geometric assumptions to be made about ventricular morphologic features, and this technique estimated systematically higher values for end-systolic volume and consequently lower estimates for stroke volume, cardiac output, and ejection fraction. We acknowledge that there were substantial differences (mean difference: >25%) between the long-axis and short-axis estimations. Despite these differences, both methods detected very high left ventricular output, without obvious alteration of ejection fraction, in the infants with a previous diagnosis of PDA. Although additional work is needed to assess the appropriate approach for volume estimations, our preliminary estimates showed test-retest values that suggested that, as in the adult population, CMRI can provide precise estimations of functional variables.21

Establishing the feasibility of CMRI for preterm infants opens a range of possibilities. CMRI provides additional techniques to define quantitative blood flow, intracardiac flow patterns, wall stress mapping, characterization of normal and damaged tissue, and molecular events in the endothelium.8 Precise quantitative data on cardiac function should also allow studies of therapeutic interventions such as inotropic or antiinflammatory treatments for cardiac dysfunction with acceptable and practical sample sizes. Therefore, CMRI has the potential to make significant contributions to research and the improvement of care for sick preterm infants. In particular, the possibility of acquiring CMRI scans without breath-holding provides the possibility of developing novel MRI sequences that optimize imaging without the time constraints of breath-holding, which may improve image quality significantly.

ACKNOWLEDGMENT
We acknowledge the ongoing work and support of Alan Groves, MD.

REFERENCES
Hydrops Fetalis: A Retrospective Review of Cases Reported to a Large National Database and Identification of Risk Factors Associated With Death

Matthew E. Abrams, MD⁎, Keith S. Meredith, MD⁎, Paula Kinnard, RN⁎, Reese H. Clark, MD⁎

⁎Center for Research and Education, Pediatrix Medical Group, Sunrise, Florida; ⁎Division of Neonatology, Phoenix Children’s Hospital, Phoenix, Arizona

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

ABSTRACT

OBJECTIVES. The objectives were (1) to identify the causes for hydrops fetalis neonates admitted for neonatal intensive care with the diagnosis of hydrops fetalis and (2) to identify the risk factors associated with death.

METHODS. A retrospective review of a large national data set was performed.

RESULTS. There were a total of 253,651 discharges from 162 NICUs in the database; 598 patients were identified with a report of hydrops fetalis. The most common associated diagnoses were congenital heart problems (13.7%), abnormalities in heart rate (10.4%), twin-to-twin transfusion (9%), congenital anomalies (8.7%), chromosomal abnormalities (7.5%), congenital viral infections (6.7%), congenital anemia (5%), and congenital chylothorax (3.2%). Of those 598 neonates, 115 were transferred either to another hospital or to another service, 215 died before discharge, and 267 were discharged from the hospital. One patient did not have a discharge type listed and was not included in the outcome analysis. Mortality rates were highest among neonates with congenital anomalies (57.7%) and lowest among neonates with congenital chylothorax (5.9%). Factors that were associated independently with death in logistic regression analyses were younger gestational age, low 5-minute Apgar score, and need for high levels of support during the first day after birth (higher levels of inspired oxygen support and more often treated with high-frequency ventilation).

CONCLUSIONS. The risk of death among neonates with hydrops fetalis depends on the underlying diagnosis and is highest for those who are born more prematurely and those who are most ill immediately after birth. Information from this large study should prove useful for planning prospective studies and providing prenatal counseling to parents with an affected fetus.
Hydrops fetalis continues to be a challenging entity in neonatal/perinatal medicine. Hydrops, an end-stage process for a number of fetal diseases, results in tissue edema and effusions of multiple body cavities. The cause of hydrops fetalis is multifactorial, and the condition is often associated with high mortality rates, despite improvements in diagnosis and management.

Because of early prenatal diagnosis and intervention, Rh disease is now a relatively uncommon cause of hydrops fetalis in neonates. Underlying mechanisms such as cardiac disease and arrhythmias are relatively straightforward to diagnose, whereas other disorders, such as lysosomal storage diseases, require careful investigation and specialty laboratories. There are many case reports in the literature attributing the underlying cause of hydrops to a number of disease entities. Our analytical approach to these data was descriptive in nature. Specific database tables within the data warehouse used for this analysis were “patients,” “admissions,” and “diagnoses.” All reports of hydrops fetalis collected within the diagnosis table were reviewed. For each patient with a diagnosis of hydrops fetalis, all other associated diagnoses were reviewed for assignment of a potential etiologic cause. In addition, data on causes of death and problems at discharge were reviewed for determination of the “primary” etiologic factor associated with hydrops fetalis. Differences in the demographic characteristics of patients who died and those who were discharged home were analyzed by comparing the 2 population samples with univariate analyses. Continuous variables (estimated gestational age and birth weight) were evaluated with 2-tailed t tests. Categorical variables (eg, race and gender) were evaluated with 2-tailed χ² tests. Nonparametric data were assessed with Kruskal-Wallis analysis of variance. After univariate analyses, we used multivariate logistic regression to calculate the adjusted odds ratio for death by comparing the neonates who died with those who were discharged home. Transferred patients were not included in this analysis. We incorporated in the logistic regression analysis the variables found in univariate analyses to be different for the treatment groups at a probability of <.1. Birth weight and gestational age were entered into the model as continuous variables. Cases with missing values for any of the independent variables were excluded from the analyses. All statistical analyses were performed by using JMP 5.0.1a (SAS Institute, Cary, NC).

RESULTS

Study Population

There were a total of 253 651 discharges from 162 NICUs within the database during the study period (Table 1). We identified 598 patients (0.23%) with a report of hydrops fetalis in the diagnosis table. Of those 598 neonates, 115 (19%) were transferred either to another hospital or to another service, 215 (36%) died before
discharge, and 267 (45%) were discharged from the hospital. One patient did not have a discharge type listed and is not included in the outcome analysis. Characteristics of the patients in each outcome group are shown in Table 1.

### Diseases Associated With Hydrops Fetalis

Of the 598 patients with a diagnosis of hydrops fetalis, a plausible cause could be found for 441 (73.7%) (Table 2). For 157 (26.3%), a cause could not be determined, and the degree of evaluation to determine a cause varied. Of those 157, 94 (60%) had reports of normal chromosomes, 92 (59%) had negative evaluation results for congenital viral infections, and 97 (61%) had reports of normal cardiac echocardiographic findings; 88 (56%) had reports of all 3. Forty-five (28.6%) of the 157 patients died within 3 days after birth, and 10 (6.4%) were transferred to another hospital at 3 days of age.

For the 441 patients with a cause, the most common diagnoses associated with hydrops fetalis were congenital infections (285 patients), maternal blood pressure disorder (119 patients), and fetal anomalies (94 patients). The most common congenital infections were listeria monocytogenes (3 patients) and cytomegalovirus (2 patients). The most common maternal blood pressure disorder was preeclampsia (9 patients). The most common fetal anomalies were congenital heart defects (7 patients) and skeletal anomalies (4 patients).
tal heart disease ($n = 82; 13.7\%$), cardiac arrhythmias ($n = 62; 10.4\%$), twin-to-twin transfusion ($n = 54; 9\%$), congenital anomalies ($n = 52; 8.7\%$), chromosomal abnormalities ($n = 45; 7.5\%$), congenital viral infections ($n = 40; 6.7\%$), congenital anemia ($n = 30; 5\%$), and congenital chylothorax ($n = 19; 3.2\%$). Five of the pa-

<table>
<thead>
<tr>
<th>TABLE 2 Causes and Outcomes</th>
<th>Associated Diagnosis</th>
<th>Distribution of Cases</th>
<th>Outcome According to Cause, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of Cases</td>
<td>Proportion of All Cases, %</td>
<td>Died</td>
</tr>
<tr>
<td>Cause of hydrops uncleara</td>
<td>157</td>
<td>26.3</td>
<td>59 (37.6)</td>
</tr>
<tr>
<td>Isomunization</td>
<td>27</td>
<td>4.5</td>
<td>6 (22.2)</td>
</tr>
<tr>
<td>Rh isomunization</td>
<td>25</td>
<td>4.2</td>
<td>6 (24)</td>
</tr>
<tr>
<td>Other</td>
<td>2h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonimmune</td>
<td>414</td>
<td>69.2</td>
<td>150 (25.1)</td>
</tr>
<tr>
<td>Congenital heart disease</td>
<td>82</td>
<td>13.7</td>
<td>32 (39)</td>
</tr>
<tr>
<td>Cardiomyopathy</td>
<td>23</td>
<td>3.8</td>
<td>7 (30.4)</td>
</tr>
<tr>
<td>Ebstein anomaly</td>
<td>14</td>
<td>2.3</td>
<td>6 (42.9)</td>
</tr>
<tr>
<td>Complex congenital heart disease (specific abnormality not reported)</td>
<td>8</td>
<td>1.3</td>
<td>2 (28.6)</td>
</tr>
<tr>
<td>Coarctation of aorta</td>
<td>7</td>
<td>1.2</td>
<td>2 (28.6)</td>
</tr>
<tr>
<td>Hypoplastic left heart</td>
<td>5</td>
<td>0.8</td>
<td>3 (60)</td>
</tr>
<tr>
<td>Hypoplastic right heart</td>
<td>5</td>
<td>0.8</td>
<td>3 (60)</td>
</tr>
<tr>
<td>Other heart defects</td>
<td>20i</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anrhythmias</td>
<td>62</td>
<td>10.4</td>
<td>14 (22.6)</td>
</tr>
<tr>
<td>Supraventricular tachycardia</td>
<td>43</td>
<td>7.2</td>
<td>8 (18.6)</td>
</tr>
<tr>
<td>Fetal arrhythmia</td>
<td>8</td>
<td>1.3</td>
<td>4 (50)</td>
</tr>
<tr>
<td>Wolff-Parkinson-White syndrome</td>
<td>6</td>
<td>1.0</td>
<td>1 (16.7)</td>
</tr>
<tr>
<td>Other heart rate anomalies</td>
<td>5g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twin-to-twin transfusion</td>
<td>54</td>
<td>9.0</td>
<td>21 (38.9)</td>
</tr>
<tr>
<td>Congenital anomalies</td>
<td>52</td>
<td>8.7</td>
<td>30 (57.7)</td>
</tr>
<tr>
<td>Syndrome not defined</td>
<td>21</td>
<td>3.5</td>
<td>10 (50)</td>
</tr>
<tr>
<td>Cystic hygroma</td>
<td>10</td>
<td>1.7</td>
<td>8 (80)</td>
</tr>
<tr>
<td>Other anomalies</td>
<td>21i</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chromosomal abnormality</td>
<td>45</td>
<td>7.5</td>
<td>16 (35.6)</td>
</tr>
<tr>
<td>Trisomy 21h</td>
<td>30</td>
<td>5.0</td>
<td>10 (33.3)</td>
</tr>
<tr>
<td>Other chromosomal anomalies</td>
<td>15h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anemia</td>
<td>30</td>
<td>5.0</td>
<td>7 (23.3)</td>
</tr>
<tr>
<td>Unknown cause for anemia (nonimmune)</td>
<td>18</td>
<td>3.0</td>
<td>5 (27.8)</td>
</tr>
<tr>
<td>Fetal-maternal transfusion</td>
<td>7</td>
<td>1.2</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Thalassemia</td>
<td>5</td>
<td>0.8</td>
<td>2 (40)</td>
</tr>
<tr>
<td>Congenital chylothorax</td>
<td>19</td>
<td>3.2</td>
<td>2 (50)</td>
</tr>
<tr>
<td>Other</td>
<td>30</td>
<td>5.0</td>
<td>13 (43.3)</td>
</tr>
<tr>
<td>Fetal demise of twin or triplet</td>
<td>5</td>
<td>0.8</td>
<td>2 (40)</td>
</tr>
<tr>
<td>Inborn error of metabolism</td>
<td>5</td>
<td>0.8</td>
<td>4 (80)</td>
</tr>
<tr>
<td>Other</td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Viral</td>
<td>40</td>
<td>6.7</td>
<td>15 (37.5)</td>
</tr>
<tr>
<td>Viral infection reported but organism not reported</td>
<td>17</td>
<td>2.8</td>
<td>6 (35.3)</td>
</tr>
<tr>
<td>Parvovirus</td>
<td>9</td>
<td>1.5</td>
<td>2 (22.2)</td>
</tr>
<tr>
<td>Cytomegalovirus</td>
<td>8</td>
<td>1.3</td>
<td>6 (75)</td>
</tr>
<tr>
<td>Other viral</td>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We limited the data reported to diagnoses with $\geq5$ patients.

a Duffy antibodies, $n = 1$; Kell antibodies, $n = 1$.

b Congenital cardiomegaly, $n = 4$; aortic valve stenosis, $n = 3$; atrioventricular canal, $n = 3$; pulmonary valve atresia, $n = 2$; pulmonary valve stenosis, $n = 2$; truncus arteriosus, $n = 2$; double-outlet right ventricle, $n = 1$; right ventricular hypertrophy, $n = 1$; single ventricle, $n = 1$; total anomalous pulmonary venous return, $n = 1$.

c Atrial flutter, $n = 4$; heart block, $n = 1$.

d Five of the patients with hydrops fetalis and a history of twin-to-twin transfusion had reports of fetal siblings who died in utero.

e Tetaroma, $n = 4$; brain abnormality/hypotonia, $n = 4$; dwarf, $n = 3$; hydrocephalus, $n = 3$; diaphragmatic hernia, $n = 2$; anencephaly, $n = 1$; cerebral venous fistula, $n = 1$; cystic adenomatoid malformation, $n = 1$; meningomyelocele, $n = 1$; vascular ring, $n = 1$.

f Some patients with trisomy 21 had additional anomalies reported (atrioventricular canal, $n = 6$; congenital anemia, $n = 2$; congenital leukemia, $n = 2$; arrhythmia, $n = 1$; cystic hygroma, $n = 1$).

g Specific abnormality not reported, $n = 6$; Turner syndrome/lysogenic hygroma, $n = 3$; trisomy 18, $n = 2$; chromosome 13 deletion, $n = 1$; XYY syndrome, $n = 1$; Klinefelter syndrome, $n = 1$; ring 18 chromosome, $n = 1$.

h Myopathy/myotonia, $n = 4$; renal abnormality/lymphomyelophthisis, $n = 1$; plasma-40, $n = 2$; mirror syndrome, $n = 2$; Noonan syndrome, $n = 2$; acardia twin/pump twin with hydrops, $n = 1$; Ballantyne syndrome, $n = 1$; congenital leukemia, $n = 1$; congenital lymph hemangiomia, $n = 1$; infantile arterial calcification, $n = 1$; pancytopenia, $n = 1$.

i Viral myocarditis, $n = 4$; congenital herpes, $n = 2$.
tients with hydrops fetalis and a history of twin-to-twin transfusion had reports of fetal siblings who died in utero.

Risk Factors Associated With Death

Univariate analyses showed that, compared with neonates who were discharged from the hospital, neonates who died were smaller and more immature (but had higher z scores, which suggests more edema) (Table 1). In addition, neonates who died were sicker in the period immediately after delivery (lower Apgar scores, higher levels of inspired oxygen, and more often treated with high-frequency ventilation) and had lower platelet counts (Table 1). The mortality rate was highest among neonates with congenital anomalies (mortality rate: 57.7%) and lowest among neonates with congenital chylothorax (mortality rate: 5.9%). Factors that were associated independently with death in logistic regression analyses were younger gestational age, low 5-minute Apgar score, and need for high levels of support during the first day after birth (higher levels of inspired oxygen support and more often treated with high-frequency ventilation).

DISCUSSION

Using a national data set, we have described the largest review of underlying causes of hydrops fetalis and identified risk factors for death. The most common diagnoses associated with hydrops fetalis were congenital heart problems, abnormalities in heart rate, twin-to-twin transfusion, congenital anomalies, chromosomal abnormalities, congenital viral infections, congenital anemia, and congenital chylothorax. Most previous studies had sample sizes of <100 patients, and only 2 of them attempted to identify risk factors associated with death.1,2,11 Similar to results reported by previous investigators,1,2,11 we noted that the prognosis for hydrops fetalis depended on the cause. In our data, the mortality rate was highest among neonates with congenital anomalies (mortality rate: 57.7%) and lowest among neonates with congenital chylothorax (mortality rate: 5.9%). Infants who died were more likely to be more premature and were sicker after birth (lower 5-minute Apgar scores, higher levels of inspired oxygen support, and more often treated with high-frequency ventilation during the first day after birth).

There are some limitations of a retrospective review of any data set that is accumulated as part of medical chart documentation. Retrospective studies are limited by incomplete data. Test or autopsy results that returned after infants died, were transferred, or were discharged might not have been entered into the database. In addition, early prenatal diagnosis may lead to early termination of the pregnancy. We did not obtain information regarding the timing and severity of prenatal presenta-

REFERENCES


COOLING CAP IS APPROVED

“A head-cooling cap that can prevent or reduce brain damage in infants starved of oxygen at birth won federal approval. The Food and Drug Administration said the device, called the Cool-Cap, could cut rates of death and disability among the estimated 5000 to 9000 such children born each year. The device, made by the Olympic Medical Corporation of Seattle, works by maintaining a flow of chilled water through a cap placed on a newborn’s head.”


Noted by JFL, MD
Cost-effectiveness of Alternative Strategies for Tuberculosis Screening Before Kindergarten Entry

Valerie J. Flaherman, MD, MPH\textsuperscript{a}, Travis C. Porco, PhD, MPH\textsuperscript{b}, Elliot Marseille, PhD, MPP\textsuperscript{c}, Sarah E. Royce, MD, MPH\textsuperscript{b}

\textsuperscript{a}Department of Pediatrics, University of California, San Francisco, California; \textsuperscript{b}California Department of Health Services, Richmond, California; \textsuperscript{c}Health Strategies International, Orinda, California

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

ABSTRACT

OBJECTIVE. We undertook a decision analysis to evaluate the economic and health effects and incremental cost-effectiveness of using targeted tuberculin skin testing, compared with universal screening or no screening, before kindergarten.

METHODS. We constructed a decision tree to determine the costs and clinical outcomes of using targeted testing compared with universal screening or no screening. Baseline estimates for input parameters were taken from the medical literature and from California health jurisdiction data. Sensitivity analyses were performed to determine plausible ranges of associated outcomes and costs. We surveyed California health jurisdictions to determine the prevalence of mandatory universal tuberculin skin testing.

RESULTS. In our base-case scenario, the cost to prevent an additional case of tuberculosis by using targeted testing, compared with no screening, was $524,897. The cost to prevent an additional case by using universal screening, compared with targeted testing, was $671,398. The incremental cost of preventing a case through screening remained above $100,000 unless the prevalence of tuberculin skin testing positivity increased to \( \geq 10\% \). More than 51\% of children entering kindergarten in California live where tuberculin skin testing is mandatory.

CONCLUSIONS. The cost to prevent a case of tuberculosis by using either universal screening or targeted testing of kindergarteners is high. If targeted testing replaced universal tuberculin skin testing in California, then $1.27 million savings per year would be generated for more cost-effective strategies to prevent tuberculosis. Improving the positive predictive value of the risk factor tool or applying it to groups with higher prevalence of latent tuberculosis would make its use more cost-effective. Universal tuberculin skin testing should be discontinued, and targeted testing should be considered only when the prevalence of risk factor positivity and the prevalence of tuberculin skin testing positivity among risk factor–positive individuals are high enough to meet acceptable thresholds for cost-effectiveness.
ROUTINE TUBERCULIN SKIN testing (TST) of children entering kindergarten is often practiced in the United States, despite recommendations of major policymaking organizations. One reason for its persistence may be that these organizations have not collaborated on specific alternative recommendations for specific ages. For example, in 2004, the Pediatric Tuberculosis Collaborative Group (PTCG) released guidelines recommending that universal TST should be replaced by risk factor screening followed by targeted TST for individuals with ≥1 positive response on the risk factor questionnaire (RQ). The recommended RQ is based on evidence from a variety of studies in pediatric populations and includes questions on country of birth, travel history, tuberculosis exposure, and close contact with someone with positive TST results. However, the PTCG did not specify optimal age groups for testing, and its guidelines suggest that testing in older age groups may be preferable.

In the absence of specific guidelines, public health officials have made individual decisions based on the available evidence and local preferences. In California, local health officers mandate universal prekindergarten TST in many counties and school districts. Universal TST is also required by the state’s Medicaid program, through the screening component of the Early and Periodic Screening Diagnosis and Treatment benefit. A decision analysis showed moderate cost-effectiveness for universal screening in Santa Clara county in 1995, and individual health officers and program officers have used these findings to make decisions regarding screening. Children in districts and jurisdictions that do not require universal TST may receive TST, risk factor screening followed by targeted TST (TT), or no screening at all at the well-child check before kindergarten entry. Similar inconsistencies have been noted in other states.

TT has been shown to be more cost-effective than universal screening for some populations of US children. However, the cost-effectiveness of using the newly developed RQ for prekindergarten screening has not been evaluated. To determine the economic and health consequences of alternative strategies for screening children in this age group, we undertook a decision analysis comparing mandatory universal TST, mandatory TT using the RQ, and no mandatory screening. We used a range of estimates for risk factor prevalence, risk of progression to disease, and proportion of TST-positive individuals in the population. We also examined the possible costs and health outcomes expected to be associated with these 3 strategies in California.

METHODS

Model
We developed a computer-based, deterministic, decision-analysis model that calculated costs and benefits from the perspective of the health care system. Cost per case of tuberculosis averted was the key outcome. Incremental costs and health outcomes were calculated by comparing results for universal screening versus TT and for TT versus no screening. Clinical and economic outcomes were discounted at a rate of 3% annually over 20 years. All costs were converted to 2004 US dollars by using the medical care component of the Consumer Price Index. The model was implemented in CLISP (Fig 1). Screening alternatives evaluated, model characteristics, and data sources used for input specification are discussed below.

Screening Strategies Evaluated
The 3 screening alternatives were as follows. The first was universal screening, in which all children would receive TST before kindergarten entry. The second alternative was risk factor screening. Before kindergarten entry, all children would receive risk factor screening for tuberculosis with the RQ recommended by the PTCG. Children with positive results for ≥1 risk factor would receive TST. The third alternative was no screening. Before kindergarten entry, TST would not be administered unless a provider suspected tuberculosis (eg, the child had symptoms of tuberculosis disease or the child was identified as a contact of a patient with an active case of tuberculosis).

We made the following clinical assumptions. (1) All children entering kindergarten are between 4.5 and 6 years of age. (2) Screening would not identify any prekindergarten children with active tuberculosis disease beyond those who would have been identified clinically. (3) Positive TST results for children 4.5 to 6 years of age would not result in a source case investigation (consistent with current national guidelines). (4) Tuberculosis disease from 5 to 10 years of age would not initiate a contact investigation. (5) Isoniazid therapy would reduce the incidence of tuberculosis by 70%.

Model Input Estimation and Sensitivity Analyses
Estimates for model inputs were obtained from the medical literature and from data provided from selected California health jurisdictions. The medical literature was searched by using Medline, manual searches of references in available literature articles, and personal questioning of tuberculosis control experts. All primary sources were obtained and examined, with questioning of authors if available.

We established a base-case scenario by using the best estimates available for each parameter. For each input, we then determined a plausible range for sensitivity analyses, on the basis of high and low estimates found in the relevant literature. We conducted 1-way sensitivity analyses for each of 4 key model inputs. We also conducted 2-way sensitivity analyses for selected pairs of inputs. The probabilities used at baseline, the minimal...
FIGURE 1
Decision tree for cost-effective analysis. Branches that represent screening options begin from square decision nodes, branches that represent chance events begin from circular chance nodes, and triangular terminal nodes represent possible outcomes. A, Alternative screening approaches. B, Decision tree for individuals who received TST. LTBI indicates latent tuberculosis infection.
and maximal values for each parameter from the medical literature, and the plausible ranges used for sensitivity analyses are presented in Table 1. The rationale for the choice of baseline values is described in the Appendix.

We used thresholds for cost-effectiveness of $50 000 and $100 000. These thresholds were chosen because $50 000 per quality-adjusted life year (QALY) saved is used frequently as an appropriate cost-effectiveness threshold for clinical and preventive interventions,23,24 and we estimated that, on the basis of expected mortality rates for our cohort, individuals who developed tuberculosis disease would lose between 1 and 2 QALYs. We also used $100 000 per QALY because there is substantial evidence that the traditional $50 000 threshold may be too low.25

Published reports of the specificity of TST range from 90% to 99%.26–29 However, this range was inconsistent with our base-case parameters. Our base-case parameter for the prevalence of TST positivity in the risk factor-negative population was 0.37%, which implies a specificity of not less than 99.6% in that group. The true specificity of TST in our population is not known. However, with higher assumed specificity, any screening intervention seems more cost-effective. Therefore, to allow comparison with current practice in the most conservative manner, we assumed as follows. For our base-case scenario, we postulated that TST had different test characteristics in these 2 populations because of different exposure to BCG vaccine and atypical mycobacteria, both of which may tend to increase false-positive rates and are more common in areas from which individuals tend to immigrate to California.30–32 We then used the highest published specificity, 99%, for the risk factor–positive population and assumed a higher specificity of 99.9% for the risk factor–negative population. We then performed a sensitivity analysis by using a high specificity of 99.9% for both populations.

Costs
Cost estimates are presented in 2004 US dollars. Major costs are summarized in Table 2. A detailed description of cost derivation is available from the corresponding author.

Cost of TST
The cost of a tuberculin skin test was assumed to be $10.21, based on 2003 Medicare Part B Extract Summary System (BESS) data.33 We made a conservative estimate that 5% must be repeated, leading to an expected cost per patient tested of $10.73. We assumed that an average of 10 minutes of licensed practical or licensed vocational nurse time would be needed to read, to interpret, and to chart the TST results, and we obtained data on estimated wage and fringe benefits from the Bureau of Labor Statistics.34 Total costs were $14.29 for each patient who had a TST both placed and read.

Cost of Risk Factor Screening
The RQ was assumed to require 2 minutes of licensed practical nurse time and 0.5 minutes of physician time for review of results, leading to a total cost of $1.41.

Cost of Evaluation and Treatment of TST-Positive Cases Identified Through Screening
For each positive TST result in the prekindergarten screening cohort, we assumed 1 medical visit and 1 chest radiograph, at costs based on BESS data.33 We assumed that all those who completed therapy would use 9 months of isoniazid, at an estimated medication cost based on 2004 Red Book wholesale prices,35 and that those who did not complete therapy would receive 2 months of isoniazid. We assumed 10 minutes of time for nurse management of refills and symptoms. Isoniazid-associated hepatitis is a very rare event in this age group. A health department-based study found no cases of hepatotoxicity among 1468 children 0 to 14 years of age,36 and no case report of significant toxicity from this cause in children <7 years of age could be found in the liter-

| TABLE 1 | Tuberculosis-Associated Probabilities From Available Literature Reports and Range for Sensitivity Analysis |
|----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Parameter | Base-Case Value | Minimal Value | Maximal Value | Range for Sensitivity Analysis |
| Probability child is risk factor positive, % | 52.87 | 20–60 |
| Probability risk factor-positive child is TST positive, % | 1.592 | 0.832 | 0.11–1 |
| Probability risk factor-negative child is TST positive, % | 0.37 | 0.16 | 0.08–0.83 | 0.11–1 |
| Risk of progression to disease, cases per 100 000 person-y | 90.237 | 41.64 | 55.734 | 30–120 |
| Probability of treatment completion given TST positivity, % | 74.4 | 58.6 | 90.5 |

| TABLE 2 | Cost Inputs for Decision Analysis |
|----------------|-----------------|-----------------|-----------------|
| Component | Cost, 2004 US Dollars |
| TST performed | 10.73 |
| TST results read | 3.56 |
| RQ | 1.41 |
| Treatment and follow-up care for positive TST results | 306.00 |
| Evaluation and treatment of tuberculosis disease, mean | 9475.00 |
Therefore, our model did not include any costs for treatment of hepatotoxicity.

Costs of Active Tuberculosis Disease
We estimated the average cost for a future case of tuberculosis, for age groups 5 to 14 years and 15 to 24 years. The components of this cost were based on expert opinion regarding the standard of care, although we note that there might be variation among practitioners. Included were costs for medical doctor visits estimated by using BESS data,\textsuperscript{33} costs for antimycobacterial drugs estimated by using the 2004 Red Book,\textsuperscript{35} costs for sputum assessments estimated by using Current Procedural Terminology codes, and costs for radiology and other laboratory studies estimated by using BESS data.\textsuperscript{33} Nurse case management time was estimated by using Bureau of Labor Statistics data. Directly observed therapy costs were calculated by averaging estimates from 3 articles.\textsuperscript{41–43} Hospitalization costs were calculated from the Centers for Disease Control and Prevention Cost of Hospitalization Study, which includes all costs of hospitalization for tuberculosis according to age group.\textsuperscript{44} Costs of contact investigation for those developing active tuberculosis in adulthood were derived from a California cohort of investigations in 1999 to 2000.\textsuperscript{45} Costs for contacts were derived by using the aforementioned estimates and data from the US Public Health Service study\textsuperscript{46} and the published literature regarding isoniazid toxicity in adults.\textsuperscript{36,47,48}

Survey
To determine the scope of possible changes in screening policy, we surveyed all 61 California health jurisdictions regarding whether TST was required for school entry at kindergarten or at any other time. We asked jurisdictions that had access to TST results from prekindergarten screening to provide summary information, including data on place of birth and incidence of TST positivity. Population estimates for California children 5 to 9 years of age were obtained from the California Department of Finance for 2004.\textsuperscript{49} The survey instrument is available from the corresponding author.

RESULTS

Base-Case Results
Table 3 shows costs and projected disease outcomes for universal testing, TT using the RQ, and no mandatory screening, with the assumptions of our base-case scenario. Net total health care system costs per 100,000 children were $61,393 for no screening, $1,029,532 for TT, and $1,531,294 for universal screening. For 100,000 individuals presenting at prekindergarten age, 6.5, 4.7, and 3.9 cases are expected to arise with no screening, TT, and universal screening, respectively.

In this base-case scenario, TT would avert 1.8 cases of tuberculosis per 100,000 children screened, compared with no screening, and universal screening would avert an additional 0.74 cases of tuberculosis per 100,000 children screened, compared with TT. The cost to prevent an additional case of tuberculosis with TT, as opposed to no screening, was $524,897. The cost to prevent an additional case of tuberculosis with universal screening, instead of TT, was $671,398.

Of 100,000 children screened, no children would begin isoniazid prophylaxis with the no-screening alternative, 795 children would begin isoniazid prophylaxis with TT, and 1,170 children would begin isoniazid prophylaxis with universal screening. Because 795 children would begin isoniazid prophylaxis with TT who would not have received isoniazid with no screening, and because TT would avert 1.8 cases of tuberculosis, compared with no screening, 440 children would be treated with isoniazid to prevent 1 case of tuberculosis if TT were used instead of no screening. Because 375 children would begin isoniazid prophylaxis with TT who would not have received isoniazid with TT, and because universal screening would avert 0.74 cases of tuberculosis, compared with TT, 507 children would begin isoniazid prophylaxis to prevent 1 additional case of tuberculosis if universal screening were used instead of TT.

Sensitivity Analyses
The incremental cost of using universal testing instead of TT was very sensitive to variations in the probability of TST positivity in the risk factor–negative population. Varying this parameter over a plausible range resulted in estimates of $316,054 to $11,804,901 per case of tuberculosis prevented by using universal testing instead of TT (Fig 2A). The incremental cost-effectiveness of using TT instead of no screening was very sensitive to variations

<table>
<thead>
<tr>
<th>TABLE 3</th>
<th>Base-Case Results for 100 000 Children Screened</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Screening</td>
</tr>
<tr>
<td>Expected tuberculosis in next 20 y, cases per 100 000</td>
<td>6,496</td>
</tr>
<tr>
<td>Expected initiation of isoniazid treatment, cases per 100 000</td>
<td>0</td>
</tr>
<tr>
<td>Program cost, $</td>
<td>0</td>
</tr>
<tr>
<td>Net discounted medical care cost for tuberculosis cases, $</td>
<td>61,393</td>
</tr>
<tr>
<td>Net total cost, $</td>
<td>61,393</td>
</tr>
<tr>
<td>Net cost per tuberculosis case averted, compared with next-least-expensive option, $</td>
<td>524,897</td>
</tr>
</tbody>
</table>

All costs were discounted at 3% per annum.
in the probability of TST positivity in the risk factor–positive population. Varying the probability of TST positivity over a plausible range resulted in estimates of $36,984 to $653,547 per case of tuberculosis prevented by using risk factor screening instead of no screening (Fig 2B).

Both the incremental cost-effectiveness of universal screening in comparison with TT and the incremental cost-effectiveness of TT in comparison with no screening were somewhat sensitive to changes in estimates of the prevalence of risk factor positivity (Fig 2C). Both estimates of incremental cost-effectiveness were very sensitive to changes in estimates of the risk of progression to disease (Fig 2D).

In a sensitivity analysis that assumed that the specificity of the test was 99.9% for both the risk factor–negative and risk factor–positive populations, TT was shown to avert 4.6 cases of tuberculosis per 100,000 children screened, compared with no screening, at a total cost of $194,185. Therefore, TT had an incremental cost-effectiveness of $204,175 per case of tuberculosis averted, compared with TT, was unchanged in this sensitivity analysis.

Threshold Analysis
The incremental cost of using universal testing instead of TT to prevent an additional case of tuberculosis did not fall below $100,000 when any plausible range for risk factor prevalence was combined with any plausible range for the prevalence of TST positivity among the risk factor–negative children. The incremental cost-effectiveness of using TT instead of no screening fell below $50,000 when the prevalence of risk factor positivity was ≥18% and the prevalence of TST positivity among the risk factor–positive children was ≥7%. The incremental cost of using TT instead of no screening fell below $100,000 when the prevalence of risk factor positivity was ≥10% and the prevalence of TST positivity among the risk factor–positive children was ≥7% (Fig 3).

Survey Results
Forty-two (69%) of 61 California health jurisdictions responded to the survey. Of those that responded, 16 (38%) required TST before kindergarten entry. The
counties that required TST before school entry included the largest health jurisdictions. Population estimates for California showed that, in 2004, ≥51% of California children overall lived in health jurisdictions that required TST before kindergarten entry and ~252,405 children entered kindergarten in areas of California that required universal TST.

**DISCUSSION**

Our results show that a majority of California children live in jurisdictions that mandate universal screening for tuberculosis before kindergarten entry and that the incremental cost of universal screening, compared with TT, for prevention of tuberculosis disease is very high. Under our base-case scenario, the incremental cost of using universal screening in place of TT is $671,398 per tuberculosis case prevented, both for the screened child and through secondary transmission. On the basis of our survey results indicating that ≥252,405 children 5 years of age live in jurisdictions that require universal testing, each year of using TT instead of currently mandated universal testing would save California $1.27 million. Each year of using TT in place of currently mandated universal testing would result in only 1.89 additional cases of tuberculosis over the subsequent 20 years. The $1.27 million in annual savings would be better invested in more cost-effective strategies to prevent tuberculosis.

Our results support strongly the recommendations of the PTCG for discontinuing universal TST of children. The incremental cost of universal screening is very high under our base-case scenario and remains quite high over a wide range of plausible circumstances. However, our results also suggest that the RQ recommended by the PTCG is not optimal for our cohort. The positive predictive value of the recommended RQ for predicting positive TST results in our base-case population is only 1.59%, and our study shows that TT remains costly, compared with no screening, when the prevalence of TST positivity in the population screened is relatively low. Foreign birth has been shown to be the strongest predictor of positive TST results for children without clinical risk factors for tuberculosis. If foreign birth were the only question used to predict risk factor status, then TT might prove to be more cost-effective than it is using the currently recommended RQ. In Santa Clara and Los Angeles, for example, ~15% to 18% of foreign-born kindergarteners have positive TST results and 7% of the kindergarteners are foreign born. Threshold analyses show that the incremental cost-effectiveness of TT, compared with no screening, falls below $100,000 when these estimates are used for the prevalence of foreign birth and the prevalence of TST positivity among foreign-born children.

Therefore, improving the risk factor screening tool may make TT more cost-effective. The choice of age group for screening may also have a strong impact on cost-effectiveness. There are 3 reasons why screening of kindergarteners may be much less cost-effective than screening of older age groups. First, the prevalence of latent tuberculosis infection is relatively low among kindergarteners, compared with older school-aged children. Our sensitivity analyses indicate that both universal screening and TT are much more cost-effective when the prevalence of TST positivity is higher. Second, the risk of transmission from subjects who develop active tuberculosis during childhood is much lower than the risk of transmission from subjects who develop tuberculosis during adolescence or later in adulthood, leading to an overall low rate of secondary transmission for kindergarteners. Our estimates of secondary transmission would be higher if our cohort were older. Third, 5 years of age marks the beginning of the “favored age,” in which children are less likely to develop tuberculosis disease. Our sensitivity analysis of the risk of progression to disease shows that incremental costs per case of tuberculosis averted are lower in populations with a higher risk of progression to disease. Our results suggest that tuberculosis screening would be more cost-effective in older age groups, in which the prevalence of TST positivity, the risk of developing active disease, and the risk of transmitting disease are higher. However, additional research is needed to determine whether increased nonadherence and hepatotoxicity in older children outweigh the benefits of later screening.

Our results differ from those obtained in Santa Clara with a screening cohort from 1992, in which TT was found to be cost-effective. In contrast to our work, the Santa Clara study used a single question regarding place of birth to determine risk status and the prevalence of TST positivity in their foreign-born kindergartners was 18%, which would improve cost-effectiveness substantially. In addition, our study used a 20-year time horizon, whereas the Santa Clara analysis used a lifetime...
There are several important limitations to our study. First, the risk of progression to disease and the specificity of TST are both parameters that are challenging to estimate for US children. However, examination of a wide range of plausible estimates for these parameters showed no incremental cost-effectiveness less than $390,000 per tuberculosis case averted. Second, our model assumes that universal TST or TT detects no cases of active tuberculosis disease in prekindergarteners and that children with active tuberculosis have symptoms that prompt their clinicians to perform TST at the time of their pre-kindergarten well-child checks. One universal TST program identified no active cases, but another detected an average of 0.2 cases per 100,000 kindergarteners screened per year. If these cases would not have been identified clinically in routine well-child checks, then the incremental cost-effectiveness of screening would be somewhat lower than we report. Third, this model did not incorporate QALYs, because of the paucity of data on QALYs associated with either tuberculosis disease in childhood or isoniazid treatment for kindergarteners.

With little economic or clinical argument to support universal prekindergarten TST, this practice has outlived its historical utility. Most, if not all, school districts in California fail to meet the threshold for cost-effective tuberculosis screening with either universal screening or TT using the RQ recommended by the PTCG. We recommend that universal TST before kindergarten entry be discontinued in all California school districts and that TT be considered only where and when the prevalence of risk factor positivity and the prevalence of TST positivity among risk factor–positive individuals are high enough to meet acceptable thresholds for cost-effectiveness. Savings should be redirected to more cost-effective methods to prevent tuberculosis. Screening with TT may be cost-effective in middle school and high school, and districts may wish to consider screening for those age groups while awaiting additional research regarding the optimal age for screening. Our evidence suggests that screening of immigrants may be very cost-effective, and additional research is needed to explore the cost-effectiveness of replacing the RQ recommended by the PTCG with a single question regarding county of birth. Screening at an older age with a tool with higher positive predictive value may yield most of the benefits of pre-kindergarten TT at a much lower cost.

**APPENDIX: CHOICE OF BASELINE VALUES FOR MODEL INPUT ESTIMATION**

The baseline estimate for the proportion of risk factor–positive children comes from Northern California Kaiser Permanente. The baseline estimate for the proportion of risk factor–positive children with positive TST results comes from the same cohort. The minimal value for this parameter comes from North Carolina, and the maximal value comes from Santa Clara and from Los Angeles. In the latter 2 cohorts, only foreign-born subjects were identified as risk factor positive. Because the positive predictive value of foreign birth for the outcome of positive TST results is higher than that of other questions on the RQ, this maximal value may be an overestimate of this parameter for the purposes of our analysis. The baseline estimate for the proportion of risk factor–negative children who have positive TST results comes from the Kaiser Permanente cohort. The minimal value comes from the South Bronx, and the maximal value comes from Los Angeles County data examining the proportion of United States-born children entering kindergarten who were found to have positive TST results. Because some United States-born children would have other risk factors and would screen risk factor positive on the RQ recommended by the PTCG, this maximal value may be an overestimate. The baseline estimate for the probability of treatment completion comes from San Diego. The maximal estimate for this parameter comes from Maryland.

The risk of progression to disease is very difficult to estimate for current US children. Available estimates identified from reviews and from examination of the literature come from historical populations, from largely adult populations, or from cohorts weighted heavily with very young children. Because routine isoniazid therapy was initiated for US populations in 1952, there is no recent US estimate for the risk of progression to disease in this age group. Our baseline estimate is the most widely cited estimate for this parameter, that is, 90.2 cases per 100,000 annually without discounting, from a Puerto Rican cohort enrolled in 1949 to 1951 and monitored for 18 to 20 years. Our minimal estimate for this parameter comes from a Hong Kong cohort. An alternative figure of 55.7 cases per 100,000 annually without discounting was found in 1924 to 1934 in Massachusetts.

**ACKNOWLEDGMENT**

This publication was made possible by grant KL2 RR024130 from the National Center for Research Resources, a component of the National Institutes of Health and the National Institutes of Health Roadmap for Medical Research.

**REFERENCES**

3. American Thoracic Society. Targeted tuberculin testing and
treatment of latent tuberculosis infection. Am J Respir Crit Care Med. 2000;161:5221–5247
11. Lobato MN, Hopewell PC. Mycobacterium tuberculosis infection after travel to or contact with visitors from countries with a high prevalence of tuberculosis. Am J Respir Crit Care Med. 1998;158:1871–1875
44. Taylor Z, Marks SM, Rios Burrows NM, Weis SE, Stricoff RL, Miller B. Causes and costs of hospitalization of tuberculosis
52. Stout JE, Nageswaran S, Tilson E. Effectiveness of a tuberculosis screening questionnaire in identifying children with latent tuberculosis infection. Presented at: annual meeting of the American Thoracic Society; May 22, 2005; San Diego, CA
53. County of Los Angeles. *Analysis of Results From the Mandated School TB Skin Test, 2004–2005, All Schools.* Los Angeles, CA: County of Los Angeles; 2005

Steven R. Nesheim, MD*, Bill G. Kapogiannis, MD||*, Minn M. Soe, MD, MPH*, Kevin M. Sullivan, PhD, MPH*, Elaine Abrams, MD**, John Farley, MD, MPH*, Paul Palumbo, MD*, Linda J. Koenig, PhD*, Marc Bulterys, MD, PhD**

Departments of *Pediatrics and ||Medicine, Division of Infectious Diseases, Emory University School of Medicine, and ||Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, Georgia; *Department of Pediatrics, Harlem Hospital Center, New York, New York; **Department of Pediatrics, University of Maryland, Baltimore, Maryland; *Department of Pediatrics, University of Medicine and Dentistry of New Jersey, Newark, New Jersey; Division of HIV/AIDS Prevention, National Center for HIV, STD, and TB Prevention, Centers for Disease Control and Prevention, Atlanta, Georgia

The authors have indicated they have no financial relationships relevant to this article to disclose. Dr Kapogiannis’ current affiliation is Pediatric, Adolescent, and Maternal AIDS Branch, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD. Dr Bulterys’ current affiliation is CDC Global AIDS Program, Lusaka, Zambia.

ABSTRACT

OBJECTIVE. We sought to determine the impact of highly active antiretroviral therapy on the incidence and prevalence of opportunistic infections in HIV-infected children.

METHODS. Children born from 1986 to 1998 were monitored until 2004 in the Perinatal AIDS Collaborative Transmission Study, sponsored by the Centers for Disease Control and Prevention. We determined the pre–highly active antiretroviral therapy and post–highly active antiretroviral therapy (before and after January 1, 1997, respectively) incidence rates of opportunistic infections among HIV-infected children and characterized the temporal decreases in percentages of CD4+ cells and the mortality rates among patients with and those without incident opportunistic infections.

RESULTS. The overall opportunistic infection incidence declined from 14.4 to 1.1 cases per 100 patient-years; statistically significant reductions were seen in the incidence of the most common opportunistic infections, including Pneumocystis jiroveci pneumonia (5.8 vs 0.3 cases per 100 patient-years), recurrent bacterial infections (4.7 vs 0.2 cases per 100 patient-years), extracocular cytomegalovirus infection (1.4 vs 0.1 cases per 100 patient-years), and disseminated nontuberculous mycobacterial infection (1.3 vs 0.2 cases per 100 patient-years). Kaplan-Meier analysis of time from birth to the first opportunistic infection illustrated more rapid acquisition of opportunistic infections by HIV-infected children born in the pre–highly active antiretroviral therapy era than by those born later. In the first 3 years of life, there was a faster decline in the percentage of CD4+ cells among children with opportunistic infections. The mortality rate was significantly higher among children with opportunistic infections.

CONCLUSIONS. Reduction in the incidence of opportunistic infections and prolongation of the time to the first opportunistic infection were noted during the post–highly active antiretroviral therapy era. Children who experienced opportunistic infections had higher mortality rates than did those who did not. Younger children (<3 years) who experienced opportunistic infections had faster declines in percentages of CD4+ T cells.
The introduction of protease inhibitors and nonnucleoside reverse transcriptase inhibitors for treatment of HIV infections made possible the use of antiretroviral combinations known as highly active antiretroviral therapy (HAART). Since the implementation of HAART became widespread, declines in mortality rates for HIV-infected adults and children have been well documented. The incidence rates of AIDS and specific opportunistic infections (OIs) have also declined. Declining OI incidence rates in children have been described but not in a birth cohort. Along with declining OI incidence rates, immunologic improvements have allowed individuals to discontinue both primary and secondary prophylactic measures for certain OIs; this approach has now become standard practice.

The Centers for Disease Control and Prevention (CDC)-sponsored Perinatal AIDS Collaborative Transmission Study (PACTS) and its successor study, Perinatal AIDS Collaborative Transmission Study-HIV Follow-up after Perinatal Exposure (PACTS-HOPE), followed from birth until 2004 a large, multicenter cohort of perinatally HIV-infected children born in the United States between 1986 and 1998. This period spans a number of major milestones in pediatric HIV/AIDS management and allowed us to examine the decline in the incidence of OIs among HIV-infected children.

Methods

Study Subjects and Design

PACTS is a CDC-sponsored, multicenter, prospective cohort study of mother-to-child HIV transmission and the natural history of HIV disease, enrolling HIV-exposed infants and their mothers in 4 US cities. The study was described in detail previously. Each site began enrollment as follows: New York City, 1986; Baltimore, 1989; Atlanta and Newark, 1990. Follow-up monitoring of mother-infant pairs was discontinued on September 30, 1999, 1 year after enrollment was terminated. HIV-infected PACTS enrollees were enrolled subsequently in a new follow-up study (PACTS-HOPE) that began in 2000. Clinical data on PACTS-HOPE enrollees were collected for the period between the 2 studies (October 1, 1999, to March 1, 2000) and then every 6 months. Follow-up monitoring continued until April 2004. Hereafter, PACTS and PACTS-HOPE enrollees collectively are referred to as the study cohort. An analysis was undertaken to determine the incidence and prevalence of HIV-related OIs and the impact of HAART on these events.

Definition of HAART Era

At each study visit, data were collected on the use of any antiretroviral medications since the previous visit. HAART was defined as the receipt of combination antiretroviral therapy that consisted of ≥2 antiretroviral medications. The regimens usually consisted of 2 nucleoside reverse transcriptase inhibitors and either a protease inhibitor or a nonnucleoside reverse transcriptase inhibitor; a few children received 3 nucleoside reverse transcriptase inhibitors. OIs that occurred during the pre-HAART and post-HAART eras were defined as those that occurred before and after January 1, 1997, respectively. The reasons for the choice of this date as the cutoff point are twofold. First, 1997 was the year in which many centers throughout the United States were initiating HAART for HIV-infected children.

Definition of OIs

OIs were defined as infectious illnesses (including recurrent bacterial infections [RBIs]) on the list of AIDS-qualifying conditions included in the revised CDC AIDS case definition.

Incidence Rate Calculation

The incidence rate of OI first was calculated for the pre-HAART and post-HAART periods by using a person-time approach and then was stratified according to age, gender, and race, to yield stratum-specific incidence rates with stratum-specific rate ratios and Mantel-Haenszel adjusted incidence rate ratios with 95% confidence intervals. Disease-specific incidence rates were calculated by censoring follow-up data after the first occurrence of that specific OI for a given patient; however, follow-up monitoring continued for that patient in other disease categories until a different OI occurred or, if no additional OI occurred, the end of follow-up monitoring was attained. Stratum-specific incidence rates were calculated by including only the first OI for a given patient, censoring the follow-up data at that time. If a zero cell resulted because of the lack of an OI, then the confi-
idence interval with exact mid-P method was used to obtain an upper confidence limit (Computer Programs for Epidemiologic Analysis).21 The follow-up times for the pre-HAART and post-HAART eras were the total numbers of months all subjects lived before and after January 1, 1997, respectively.

Time to First OI
Survival analysis was used to compare the time until the development of the first OI among HIV-infected children born before or after January 1, 1997. For this analysis, subjects born in the pre-HAART era who were monitored beyond January 1, 1997, had their follow-up times censored after January 1, 1997. Kaplan-Meier and adjusted survival plots controlling for gender and race were constructed for comparisons between the 2 groups described previously. To determine the effect of the post-HAART era on the time to development of a first OI, we used Cox proportional-hazard analysis to yield the final model, after testing interaction terms and assessing confounding by other covariates such as gender and race. When constructing the models, we applied multicollinearity and regression diagnostic tests to evaluate model fit.

Immunologic and Virologic Status
Immunologic status was assessed as the percentage of CD4+ cells, and virologic status was assessed through plasma HIV RNA quantification (HIV viral load). Percentage of CD4+ cells was used instead of the absolute CD4+ cell count for the analysis because the former is generally accepted as a more stable parameter, particularly during the early years of life, when there can be large fluctuations in normal absolute CD4+ cell count ranges.22 To determine the trends for these parameters, their distributions according to age were estimated through linear regression analysis with 95% confidence intervals, as described previously by Denny et al.22 Data chosen for this analysis were limited to children 0 through 72 months of age; this interval was selected because it included nearly all of the OI occurrences (195 of 211 cases; 92%). Data from the same patient at different ages were identified, and a single result was selected randomly for analysis; this approach was used to avoid overrepresentation of individual data, which might be highly correlated. The comparison of the slopes of the 2 regression lines (children with OIs and children without OIs) was analyzed through assessment of an interaction term in the linear regression model.

Mortality Analysis
Deaths of HIV-infected children were analyzed through survival analysis, to compare the time until death between children with and without OIs. Subjects born in the pre-HAART era who were monitored beyond January 1, 1997, had their follow-up times censored after January 1, 1997. Cox proportional-hazard analysis was also performed, to explore the contribution of other covariates, such as gender, race, and clinical sites of care.

Statistical Analyses
Demographic data were analyzed with SAS 9.0 (SAS Institute, Cary, NC) and Epi Info 3.3.2 (CDC, Atlanta, GA) software. Estimates of incidence density rates and rate ratios were computed by using OpenEpi software (OpenEpi, Atlanta, Georgia). Survival analyses, including Cox proportional-hazard models, were conducted with SAS software. All significance tests were 2-tailed, and a P value of ≤.05 was considered statistically significant. Confounding by a variable was defined at the analytical stage if the adjusted incidence rate ratio derived from stratification of that variable differed from the crude rate ratio by ≥5%.

RESULTS
During 1986 to 1998, 364 HIV-infected children were enrolled in PACTS (Table 1); 165 OIs occurred in 113 children (31%) during the entire period in which the study cohort was observed. One hundred fifty-three OIs occurred (in 1061 person-years of observation) during the pre-HAART era, and 12 OIs occurred (in 1101 person-years of observation) during the post-HAART era, which yielded an incidence rate ratio of 0.08 (95% con-

| TABLE 1 Characteristics of HIV-Infected Children Who Developed OIs (N = 165) in the Pre-HAART Era (Before January 1, 1997) and the Post-HAART Era |
|---|---|---|
| Gender | Pre-HAART Era (n = 153) | Post-HAART Era (n = 12) |
| Female | 78 (51.0) | 9 (75.0) |
| Male | 75 (49.0) | 3 (25.0) |
| Race | | |
| Black, non-Hispanic or not specified | 107 (69.9) | 9 (75.0) |
| White, non-Hispanic or not specified | 11 (7.2) | 1 (8.3) |
| Hispanic | 32 (20.9) | 2 (16.7) |
| Other | 3 (2.0) | 0 (0.0) |
| Calendar year of birth | | |
| 1986–1988 | 23 (15.0) | 1 (8.3) |
| 1989–1991 | 73 (47.7) | 4 (33.3) |
| 1992–1994 | 47 (30.7) | 4 (33.3) |
| 1995–1996 | 10 (6.5) | 2 (16.7) |
| 1997–1999 | 0 (0.0) | 1 (8.3) |
| Age at diagnosis of OI, y | | |
| 0–3 | 133 (86.9) | 3 (25.0) |
| 4–6 | 17 (11.1) | 3 (25.0) |
| 7–9 | 3 (2.0) | 4 (33.3) |
| 10 | 0 (0.0) | 2 (16.7) |

Total percentages may not be exactly 100% because of rounding. The unit of analysis is OI event, not individual patient. Repetition of characteristics is attributable to multiple OI events for 1 person on the same date or a different date.
The confidence interval: 0.04–0.14). The outcome for each child enrolled depended on many dynamic exposure variables over the 18-year course of the study, including age, birth date, and length of exposure in defined time periods at high-risk ages (before HAART, before trimethoprim/sulfamethoxazole, and before intravenously administered immunoglobulin) (Fig 1).

Incidence rates and the relative frequencies of specific OIs in each era are shown in (Table 2). *Pneumocystis jiroveci* pneumonia (PCP), RBIs, extraocular cytomegalovirus (CMV), and nontuberculous mycobacteria (NTM) (including *Mycobacterium avium* complex and *Mycobacterium kansasii*) were the most frequent OIs in the pre-HAART era, accounting for 37.3%, 29.4%, 9.8%, and 9.2% of OIs, respectively. In the post-HAART era, *Candida* esophagitis was the most frequent OI, accounting for 33.3% (compared with 6.5% in the pre-HAART era), with PCP (25%), RBIs (16.7%), NTM (16.7%), and extraocular CMV (9.1%) remaining the next most frequent OIs. Several OIs that had been relatively infrequent in the pre-HAART era (respiratory candidiasis, extrapulmonary cryptococcosis, cryptosporidiosis, CMV

![Diagram](image-url)
retinitis, herpes simplex virus infection, and extrapulmonary tuberculosis) had no cases in the post-HAART era. Statistically significant reductions in incidence were seen for PCP, RBIs, extraocular CMV, disseminated NTM, and respiratory candidiasis, whereas declines involving the remaining, less-frequent OIs did not achieve statistical significance (Table 2). When analysis was restricted to OIs that occurred at ≤6 years of age, there was a statistically significant decline in the incidence of Candida esophagitis (data not shown). Stated differently, in the pre-HAART era, all Candida esophagitis cases occurred in children ≤6 years of age; in the post-HAART era, 6 of 7 Candida esophagitis cases occurred in children >6 years of age. For PCP, RBIs, extraocular CMV, disseminated NTM, and respiratory candidiasis, the incidence rate ratio ranged from 0.0 to 0.14, indicating an 86%–100% reduction in the incidence of those OIs.

Because most OIs occurred in children <6 years of age, the times from birth to the first OI for all children <6 years of age who were born before versus after January 1, 1997, were compared through Kaplan-Meier analysis (Fig 2). The results of this analysis illustrated dramatically more rapid acquisition of OIs by HIV-infected children born in the pre-HAART era, compared with those born later (P < .0001, log-rank test). Only 1 OI occurred in a child born in the post-HAART era.

Age-related incidence rates were calculated for age strata of <6 years, on the basis of the same rationale.

### TABLE 2
Incidence Rates of OIs According to Anatomic Sites Among HIV-Infected Children (N = 364) in the Pre-HAART Era (Before January 1, 1997) and the Post-HAART Era

<table>
<thead>
<tr>
<th>OI and Anatomic Site</th>
<th>Pre-HAART Era</th>
<th>Post-HAART Era</th>
<th>Incidence Rate Ratio, Median (95% CI)*a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Relative Frequency, %</td>
<td>Incidence Rate, Median (95% CI), Cases per 100 Person-y</td>
</tr>
<tr>
<td>PCP</td>
<td>57</td>
<td>37.3</td>
<td>5.8 (4.4–7.5)</td>
</tr>
<tr>
<td>RBIs</td>
<td>45</td>
<td>29.4</td>
<td>4.7 (3.4–6.3)</td>
</tr>
<tr>
<td>CMV disease, extraocular</td>
<td>15</td>
<td>9.8</td>
<td>1.4 (0.8–2.4)</td>
</tr>
<tr>
<td>Disseminated NTM</td>
<td>14</td>
<td>9.2</td>
<td>1.3 (0.7–2.3)</td>
</tr>
<tr>
<td>Candidiasis, esophageal</td>
<td>10</td>
<td>6.5</td>
<td>0.9 (0.5–1.7)</td>
</tr>
<tr>
<td>Candidiasis, respiratory</td>
<td>5</td>
<td>3.3</td>
<td>0.5 (0.2–1.1)</td>
</tr>
<tr>
<td>Herpes simplex virus</td>
<td>2</td>
<td>1.3</td>
<td>0.2 (&lt;0.1–0.7)</td>
</tr>
<tr>
<td>Cryptosporidiosis, chronic intestinal</td>
<td>2</td>
<td>1.3</td>
<td>0.2 (&lt;0.1–0.7)</td>
</tr>
<tr>
<td>Cryptococcosis, extrapulmonary</td>
<td>1</td>
<td>0.7</td>
<td>0.1 (&lt;0.1–0.5)</td>
</tr>
<tr>
<td>CMV retinitis</td>
<td>1</td>
<td>0.7</td>
<td>0.1 (&lt;0.1–0.5)</td>
</tr>
<tr>
<td>Mycobacterium tuberculosis, extrapulmonary</td>
<td>1</td>
<td>0.7</td>
<td>0.1 (&lt;0.1–0.5)</td>
</tr>
<tr>
<td>Total</td>
<td>153</td>
<td>14.4 (12.2–16.9)</td>
<td>12</td>
</tr>
</tbody>
</table>

CI indicates confidence interval.

a The overall crude rate ratio is the rate in the post-HAART era divided by the rate in the pre-HAART era.

b Significant at α = .05.

FIGURE 2
Kaplan-Meier analysis of time to development of the first OI in children ≤6 years of age, comparing children who were born before January 1, 1997 (pre-HAART era), with those born after January 1, 1997 (post-HAART era).
already provided. In that age group, there were 107 first OI occurrences, yielding 103 and 4 first OIs in the pre-HAART and post-HAART eras, respectively. OI rates in the pre-HAART and post-HAART eras were 12.5 cases per 100 person-years and 0.8 cases per 100 person-years, respectively (Table 3). When the data were stratified according to age intervals of 2 years, OI incidence rates declined significantly in the 0- to 2-year and 2- to 4-year age strata, from the pre-HAART era to the post-HAART era. A similar trend was noted for the 4- to 6-year age stratum but failed to achieve statistical significance, probably because of the small number of events in those cells. Significant declines in OIs from the pre-HAART era to the post-HAART era were also noted in both gender and race groups.

In the characterization of other mediating factors, such as immunologic status, that might have contributed to the decline in OI incidence, the method of Denny et al. was adapted to analyze differences in the temporal trends of percentages of CD4+ cells and HIV RNA viral loads among patients who developed OIs, compared with those who never had an OI. Because the large majority of OIs occurred by 6 years of age, analyses were limited to percentage of CD4+ cells values from that age range and from several age strata within that range. At <3 years of age, children who never developed an OI experienced a decline in their percentage of CD4+ cells \((r = –0.28; \ P < .001, \ \text{Pearson’s test})\), as did those who had ≥1 OI \((r = –0.49; \ P < .001, \ \text{Pearson’s test})\); however, the rate of decline of percentage of CD4+ cells was slower in the former group \((P = .02)\) (Fig 3). When all children <6 years of age were analyzed, there was no significant difference in the trends in percentages of CD4+ cells among children with and without OIs \((P = .38)\). Similar analysis of the HIV RNA viral loads did not reveal any significant differences between trends \((P > .05)\).

Mortality rates among children with OIs were assessed through Kaplan-Meier analysis; results of this analysis revealed significantly higher mortality rate, with a shorter time to death, among children with OIs than among those without OIs \((P < .0001, \ \text{log-rank test})\) (Fig 4). With the Cox proportional-hazards model, the risk of death was higher among children with OIs (hazard ratio: 5.6; 95% confidence interval: 3.6–8.8), after adjustment for gender, race, and site of care.

**DISCUSSION**

In this 18-year, multicenter, prospective cohort study, there were 86%–100% reductions in the incidence rates of PCP, RBIs, extracellular CMV, disseminated NTM, and respiratory candidiasis, in a comparison of the time periods before and after January 1, 1997 (the pre-HAART and post-HAART eras, respectively). Furthermore, there was a significant prolongation in the time to the development of a first OI among children born during the post-HAART era. We explored the basis for the declining incidences by comparing the trends in percentages of CD4+ cells during the first 3 years of life for children with and without OIs, and we found a more-marked decline in the percentage of CD4+ cells among children with a history of OIs. This trend was not found among those >3 years of age, likely because >80% (136 of 165) of all OIs occurred in children <3 years of age, and the median time to the first OI was 9.3 months. Not surprisingly, the occurrence of an OI was associated significantly with an increased mortality rate.

Among the strengths of this cohort were its follow-up period and its size. Between 1986 and 1998, >2000 mother-infant pairs were enrolled and 364 HIV-infected

| TABLE 3 | Incidence Rates of First OIs Among HIV-Infected Children \((N = 364)\) in the First 6 Years of Life in the Pre-HAART Era (Before January 1, 1997) and the Post-HAART Era |
| --- | --- | --- |
| **Pre-HAART Era** | **Post-HAART Era** | **Incidence Rate Ratio, Median (95% CI)*** |
| **Age at diagnosis, y** | **n** | **Person-Time, y** | **Incidence Rate, Median (95% CI), Cases per 100 Person-y** | **n** | **Person-Time, y** | **Incidence Rate, Median (95% CI), Cases per 100 Person-y** | **Incidence Rate Ratio** |
| 0 to <2 | 85 | 413.6 | 20.6 (16.4–25.4) | 3 | 270.8 | 1.1 (0.2–3.2) | 0.05 (0.01–0.15) |
| 2 to <4 | 13 | 274.2 | 4.7 (2.5–8.1) | 0 | 145.3 | 0.0 (0.0–2.5) | 0.00 (0.00–0.49) |
| 4 to <6 | 5 | 135.0 | 3.7 (1.2–8.6) | 1 | 67.9 | 1.5 (0.02–8.2) | 0.40 (0.02–2.87) |
| **Gender** | | | | | | | |
| Male | 51 | 387.0 | 13.2 (9.8–1.7) | 1 | 166.7 | 0.6 (<0.01–3.3) | 0.05 (<0.001–0.23) |
| Female | 52 | 435.8 | 11.9 (8.9–15.7) | 3 | 317.3 | 0.0 (0.0–2.8) | 0.08 (0.02–0.23) |
| **Race** | | | | | | | |
| Black, non-Hispanic | 74 | 563.5 | 13.1 (10.3–16.5) | 4 | 330.1 | 1.2 (0.3–3.1) | 0.09 (0.03–0.23) |
| Other | 29 | 259.3 | 11.2 (7.5–16.1) | 0 | 153.9 | 0.0 (0.0–2.4) | 0.00 (0.00–0.18) |
| Overall crude | 103 | 822.8 | 12.5 (10.2–15.2) | 4 | 484.0 | 0.8 (0.2–2.1) | 0.07 (0.02–0.16) |

CI indicates confidence interval.

* The overall crude rate ratio is the rate in the post-HAART era divided by the rate in the pre-HAART era.

b Significant at \(\alpha = .05\).
children were identified; with follow-up monitoring continuing into 2004, this cohort constitutes one of the largest of its kind in the world. For the study, we were able to ascertain clinical data comprehensively because (1) prenatal HIV testing was routine at participating institutions at an early time, (2) large proportions of the HIV-exposed mother-infant pairs were enrolled at the participating sites, (3) clinical care of the patients was performed by the study investigators, and (4) clinical diagnoses were recorded by study staff members at the time of diagnosis.

Although this study was prospective, one limitation is that it was not controlled and therefore was not designed to determine causal relationships between variables and outcomes. The duration of the study, spanning many of the management milestones of the HIV/AIDS epidemic, made analysis of the data very challenging. Primary prophylaxis of OIs,23,24 graduated antiretroviral therapy uptake (monotherapy then dual therapy then HAART), intravenously administered immunoglobulin prophylaxis of bacterial infections,25–27 monitoring of therapy with quantitative HIV RNA viral load determinations,
routine voluntary HIV antibody testing for pregnant women, and antiretroviral prophylaxis of mother–to–child transmission are prime examples of these milestones. As a result of HIV testing and antiretroviral prophylaxis recommendations made in late 1994, the number of HIV-infected newborns decreased strikingly. After that time, there was an attendant decrease in the person-years of observation spent during the first 2 years of life, a time when the 2 most common OIs of HIV-infected children (PCP and RBIs) occur frequently. HIV-infected children in our cohort were born at varying times relative to the aforementioned milestones and thus differed greatly in the benefit they might have received from any particular innovation. These milestones are contextualized with the numbers of births and OIs per year, as well as with opportunity of an age group to spend person-time in any given period (Fig 1). Interestingly, the distribution of OIs followed closely the distribution of births, that is, the peak incidence of OIs was substantially older. As in the PSD study, we counted only first-time OIs. Our post-HAART data and those of the PACTG study show similar OI incidence rates, with the exception again being PCP (0.09 cases per 100 patient-years in the PACTG study and 0.3 cases per 100 patient-years in PACTS).

CONCLUSIONS

In our cohort, declining incidence rates were seen for all OIs. In some cases, the low frequency of certain OIs did not allow declining trends to achieve statistical significance, although all of these OIs seemed to decrease. These changes are attributable largely to the widespread use of HAART but also to improvements in care, such as improved identification of at-risk infants (through testing of pregnant women) and improved use of OI prophylaxis. Parallel to improvements in care has been the remarkable decrease in mother-to-child HIV transmission. The unique aspects of this long-term study allowed us to measure and to contextualize these changes. It remains to be seen whether these declining trends in OI incidence will be sustained as more children proceed through adolescence and whether a similar pattern of decline in OI incidence will be seen as HAART becomes more available to resource-poor areas.

ACKNOWLEDGMENTS

PACTS and PACTS-HOPE were funded by the CDC through cooperative agreements U64/CCU207228 (Medical and Health Research Association of New York City), U64/CCU202219 (University of Medicine and Dentistry of New Jersey, New Jersey Medical School), U64/CCU306825 (University of Maryland School of Medicine), and U64/CCU404456 (Emory University School of Medicine).

At Emory University, Vickie Grimes served as research nurse for the duration of the study, and her work is greatly appreciated. At the University of Maryland, Drs Peter Vink and Vicki Tepper served as site principal investigators, Sue Hines as study coordinator, and Katie Peery as research assistant. At the University of Medicine and Dentistry of New Jersey, Linda Bettica, RN, served as research coordinator and Jeffrey Swerdlow as database manager. We also thank the participants in the New York City PACTS Group, as follows: Bronx Lebanon Hospital, Saroj Bakshi, Genevieve Lambert, Elizabeth
Adams, and Delia Grant; Harlem Hospital Center, Susan Champion, Julia Floyd, Cynthia Freeland, Margaret Hegarty, Pamela Prince, Desiree Minnott, and Aretha Bellmore; Jacobi Hospital Center, Joanna Dobroszycki, Adell Harris, and Andrew Wiznia; Metropolitan Hospital Center, Mahrurk Banaji, Grace Canillas, Lynn Jackson, and Nancy Cruz; Medical and Health Research Association of New York City, Tina Allford, Rosalind Carter, Mary Ann Chiasson, Eileen Rillamas-Sun, Donald Thea, and Jeremy Weeden; Montefiore Medical Center, Ellie Schoenbaum and Marcelle Naccarato. Finally, at the CDC, the following staff members contributed substantially to PACTS: Martha Rogers, Nathan Shaffer, Rick Steketee, and RJ Simonds; to PACTS-HOPE: Darcy Freedman, Jeff Wiener, Bob Yang, and April Bell.

REFERENCES


---

VIOLENT JUSTICE: ADULT SYSTEM FAILS YOUNG OFFENDERS

“State laws that send some individuals under age 18 to trial and prison as adults have achieved the opposite of what the policy’s proponents intended, a new research review concludes. Transferring young people into adult systems yields substantially higher rates of later serious crimes compared with youths handled by juvenile-justice systems. Moreover, there’s no evidence that shifting some young offenders to the adult justice system prevents or reduces violence in the general population of children and teenagers. These findings come from the 14-member Task Force on Community Preventive Services, an independent group funded by federal and private sources. It’s reviewing the effectiveness of various efforts to lessen violence committed by and against youths. The task force reports that young offenders transferred to the adult system are later arrested for violent and other crimes 34 percent more frequently than are their peers sent to juvenile courts and facilities. The task force compared juveniles charged with comparable offenses. Its report appears in a supplement to the April American Journal of Preventive Medicine.”

Science News. April 21, 2007

Noted by JFL, MD
Is Victimization From Bullying Associated With Medicine Use Among Adolescents? A Nationally Representative Cross-sectional Survey in Denmark

Pernille Due, MDa,b, Ebba Holme Hansen, MScb,c, Juan Merlo, PhDd, Anette Andersen, PHDb, b, Bjørn E. Holstein, MagScientSoc a,b

aDepartment of Social Medicine, Institute of Public Health, Faculty of Health Sciences, and bDepartment of Pharmacology and Pharmacotherapy, Section of Social Pharmacy, Faculty of Pharmaceutical Sciences, University of Copenhagen, Copenhagen, Denmark; cDepartment of Clinical Sciences, Faculty of Medicine, Lund University, Malmö, Sweden

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

ABSTRACT

OBJECTIVE. The goal was to examine whether being a victim of bullying was associated with medicine use, taking into account the increased prevalence of physical and psychological symptoms.

METHODS. The study population included all students in grades 5, 7, and 9 (mean ages: 11.6, 13.6, and 15.6 years, respectively) in a random sample of schools in Denmark (participation rate: 88.5%; \( N = 5205 \)). The students reported health problems, medicine use, bullying, and a range of psychosocial conditions in an anonymous standardized questionnaire. The outcome measure was self-reported medicine use for headache, stomachache, difficulties in getting to sleep, and nervousness. The determinant was frequency of exposure to bullying, measured with 1 item.

RESULTS. In multivariate models adjusted for age and social class, we found that adolescent victims of bullying used medicine for pains and psychological problems more often than did adolescents who were not bullied. The increased odds of using medicine were not explained by the higher prevalence of symptoms among the bullied children.

CONCLUSIONS. We found victimization from bullying to be associated with medicine use, even when we controlled for the higher prevalence of symptoms among bullied victims. The medications that adolescents use can have adverse effects, in addition to the potentially health-damaging effects of bullying. Policy makers, health care professionals, and school staff should be aware that the adolescent victims of bullying are prone to excess use of medicine, and preventive actions should be taken to decrease the level of bullying as well as the use of medicine among adolescents.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-1481
doi:10.1542/peds.2006-1481

Key Words: bullying, medicine use, adolescents, population-based study

Abbreviations: HBSC—Health Behavior in School-aged Children
MOR—median odds ratio
OR—odds ratio
CI—confidence interval

Accepted for publication Mar 1, 2007
Address correspondence to Pernille Due, MD, Department of Social Medicine, Institute of Public Health, Faculty of Health Sciences, University of Copenhagen, Øster Farimagsgade 5, PO Box 2099, DK-1014 Copenhagen, Denmark. E-mail: p.due@pubhealth.ku.dk

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
Using medicine to alleviate common ailments is a frequent behavior among adolescents.1,2 The use of pain-relieving medicine and psychotropic drugs among adolescents is an important issue for health care professionals, because the substances are potentially toxic and may have harmful adverse effects.3

The patterns of medicine use according to gender and age are consistent across countries, despite considerable national differences in the prevalence of medication use. In general, girls seem to use medicines for headaches and stomachaches more often than do boys, and both prevalence and gender differences in use increase from 11 years to 15 years of age. The prevalence of medicine use for psychological problems is greater among younger boys, but the prevalence decreases with age for both genders. By the age of 15 years, the prevalence of the use of these medications is also higher among girls.1

One important stressor in adolescent life is frequent bullying,5,7 which, in turn, is clearly related to aches and psychological problems.5–18 Medicine use is a behavior that, over and above formal therapeutic indications, may reflect a general coping strategy to overcome daily stressors. Therefore, frequent bullying may increase the use of pain-relieving medicine and psychotropic drugs.

It is known that health and several health-related behaviors in adults seem to be rooted in childhood and adolescence. With a long-term perspective, early use of medicines to alleviate pain and psychological problems may prevent the learning of more appropriate ways of coping. In addition to the health-damaging effects of bullying, this practice adds the potential risk of lifelong medicine use. Despite the apparent relevance, however, we were unable to identify any previous studies dealing with the relationship between bullying and medicine use among adolescents. This study examined the association between victimization from bullying and medicine use for treatment of pain and psychological problems in a representative sample of adolescents.

METHODS

Design and Study Population
We used data from the Danish contribution to the international cross-sectional study Health Behavior in School-aged Children (HBSC) in 1998.4 The study population comprised all pupils in grades 5, 7, and 9 (average age: 11.6, 13.6, and 15.6 years, respectively) in a random sample of schools in Denmark. The general purpose of the HBSC study was to gain insight into young people’s health and health-related behaviors.

The participation of the schools in the survey relied on the previous informed consent of the headmaster, the school board, and the board of pupils at each school. Fifty-five of the 68 randomly sampled schools agreed to participate in the study. In total, 99% of the pupils present on the day of data collection participated in the survey, corresponding to 88.5% of the pupils enrolled in the relevant classes (N = 5205). Because the study was conducted anonymously, we were unable to analyze the characteristics of nonparticipants.

Data Collection
Data were collected in the classroom by means of a carefully tested questionnaire, with standardized instructions from the teacher. The pupils returned the completed questionnaires to their teacher in sealed envelopes to protect anonymity, and no school staff members had access to the completed questionnaires.

Assessment of Variables
The dependent variable “medicine use” was measured with the following item: “During the past month, did you take any tablets or medicine for (1) headache, (2) stomachache, (3) difficulties in getting to sleep, or (4) nervousness?” The responses were categorized as (1) no or (2) yes, once or yes, several times. The items have been tested to show high external validity, and qualitative analyses have shown good face and content validity of the items.19

The independent variable “bullying” was measured with the following item: “How often have you been bullied in school this term?” The responses were categorized into 3 levels, namely, (1) not at all, (2) once or twice/a few times, and (3) once per week/several times per week. Previous studies have demonstrated a high degree of validity in the applied measurement of bullying.20,21

Family social class was derived from 2 items on father’s and mother’s occupations. Children’s reports of their parents’ occupations have been shown to be valid.22 In accordance with the standards of the Danish National Institute of Social Research,23 pupils’ responses were coded from I (high) to V (low), with VI indicating parents receiving social benefits. Approximately 12% of the pupils did not provide sufficient information to allow coding of social class, and we included lack of information on family social class as a specific category in the analyses. Each pupil was coded in accordance with the parent with the highest rank. The pupils were categorized into 4 groups, namely, (1) I/II, (2) III/IV, (3) V/VI, and (4) missing data on social class.

Symptom prevalence was measured with 4 items from the validated HBSC Symptom Check List,24 as follows: “During the past 6 months, how often have you experienced (1) headache, (2) stomachache, (3) difficulties in getting to sleep, or (4) nervousness?” The responses were recoded into 3 levels, namely, (1) almost daily/more than once per week, (2) about once per week/about once per month, or (3) less often or never. Category 3 was used as reference in the comparisons.
Statistical Procedures

We used SAS 8.2 software (SAS Institute, Cary, NC) for all analyses. We performed preliminary separate logistic regression analyses for each kind of medicine use, with bullying as the independent variable, and supplemented each of these analyses with multilevel logistic regression analyses (Tables 1 and 2). Analyses including both genders were performed to test the interaction of gender with the association between bullying and medicine use. The interaction term was not significant for medicine use to alleviate headache, stomachache, or nervousness ($P = .111$, $.396$, and $.205$, respectively), whereas the interaction was significant ($P = .046$) for medicine for sleeping difficulties. Therefore, all analyses were conducted separately for boys and girls, and we present the results according to gender specifically (Tables 1 and 2).

First, we tested the crude association between bullying and medicine use for each of the symptoms separately (model 1). In model 2, age group and family social class were included in the analyses, because these 2 factors have been reported to be associated with the exposure variable bullying, as well as with the outcome variable medicine use.1,2,25,26 In model 3, the analyses were also adjusted for the prevalence of the symptom for which the medicine was taken. Associations were reported as odds ratios (ORs) with 95% confidence intervals (CIs).

We used the median OR (MOR) to express school and school class variance in medicine use.27–29 The MOR quantifies differences (ie, variance $\sigma^2$) between schools and between school classes by comparing 2 children with the same covariates but from 2 randomly chosen, different schools or school classes. This procedure yields a distribution of ORs, with 1 OR for each comparison pair. The MOR is the median of this distribution of pairwise ORs. That is, the MOR expresses how much (in median) the individual probability of using a certain medicine would increase if a child moved to another school or school class with higher use of medicine. The MOR measure is always $\geq 1$. If the MOR is 1, then there are no differences in medicine use between schools (no second-level variation). If there are considerable school differences, then the MOR is large. The measure is directly comparable to fixed-effects ORs, which makes quantification of school variance easier to appreciate in terms of the familiar ORs.

In our investigation, we used the school class-level variance to compute the school class MOR. The interpretation of this measure would be how much (in median) the individual probability of using a certain medicine would increase if an individual within a school moved to another school class with higher use of medicine. An equivalent interpretation is valid for MOR at the school level. Obviously, when a child changes schools, he or she also changes school classes. Therefore, we could calculate the school-level MOR by using the

### Table 1

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not bullied</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Bullied monthly</td>
<td>1.48 (1.2–1.8)</td>
<td>1.50 (1.3–1.8)</td>
<td>1.28 (1.0–1.6)</td>
<td>1.28 (1.0–1.7)</td>
<td>1.16 (0.9–1.4)</td>
<td>1.19 (0.9–1.4)</td>
<td>2.08 (1.5–2.9)</td>
<td>2.11 (1.6–2.6)</td>
<td>2.08 (1.5–2.9)</td>
<td>2.03 (1.4–2.9)</td>
<td>1.87 (1.3–2.5)</td>
<td>1.87 (1.3–2.5)</td>
</tr>
<tr>
<td>Bullied at least weekly</td>
<td>2.26 (1.5–3.0)</td>
<td>2.28 (1.7–3.1)</td>
<td>1.73 (1.3–2.4)</td>
<td>1.73 (1.3–2.4)</td>
<td>1.60 (1.0–2.4)</td>
<td>1.58 (1.0–2.3)</td>
<td>3.60 (2.3–5.7)</td>
<td>3.53 (2.3–5.7)</td>
<td>3.49 (2.3–5.7)</td>
<td>3.46 (2.3–5.7)</td>
<td>3.44 (2.3–5.7)</td>
<td>3.44 (2.3–5.7)</td>
</tr>
<tr>
<td>5th grade (11-y-old students)</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>7th grade (13-y-old students)</td>
<td>1.00 (0.9–1.3)</td>
<td>1.00 (0.9–1.3)</td>
<td>0.60 (0.4–0.9)</td>
<td>0.60 (0.4–0.9)</td>
<td>0.60 (0.4–0.9)</td>
<td>0.60 (0.4–0.9)</td>
<td>0.35 (0.1–0.6)</td>
<td>0.35 (0.1–0.6)</td>
<td>0.35 (0.1–0.6)</td>
<td>0.35 (0.1–0.6)</td>
<td>0.35 (0.1–0.6)</td>
<td>0.35 (0.1–0.6)</td>
</tr>
<tr>
<td>9th grade (15-y-old students)</td>
<td>1.23 (0.9–1.5)</td>
<td>1.23 (0.9–1.5)</td>
<td>0.31 (0.2–0.6)</td>
<td>0.31 (0.2–0.6)</td>
<td>0.31 (0.2–0.6)</td>
<td>0.31 (0.2–0.6)</td>
<td>0.20 (0.1–0.5)</td>
<td>0.20 (0.1–0.5)</td>
<td>0.20 (0.1–0.5)</td>
<td>0.20 (0.1–0.5)</td>
<td>0.20 (0.1–0.5)</td>
<td>0.20 (0.1–0.5)</td>
</tr>
<tr>
<td>Family social class I/II (high)</td>
<td>1.13 (1.0–1.3)</td>
<td>1.13 (1.0–1.3)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
</tr>
<tr>
<td>Family social class III/IV</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
<td>1.00 (0.9–1.1)</td>
</tr>
<tr>
<td>Symptom never or seldom</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Symptom monthly</td>
<td>5.32 (4.4–6.4)</td>
<td>5.32 (4.4–6.4)</td>
<td>4.71 (3.5–6.7)</td>
<td>4.71 (3.5–6.7)</td>
<td>4.71 (3.5–6.7)</td>
<td>4.71 (3.5–6.7)</td>
<td>4.71 (3.5–6.7)</td>
<td>4.71 (3.5–6.7)</td>
<td>4.71 (3.5–6.7)</td>
<td>4.71 (3.5–6.7)</td>
<td>4.71 (3.5–6.7)</td>
<td>4.71 (3.5–6.7)</td>
</tr>
<tr>
<td>MOR for class</td>
<td>1.34</td>
<td>1.00</td>
<td>3.69</td>
<td>1.00</td>
<td>1.34</td>
<td>3.69</td>
<td>1.00</td>
<td>1.34</td>
<td>3.69</td>
<td>1.00</td>
<td>1.34</td>
<td>3.69</td>
</tr>
<tr>
<td>MOR for school</td>
<td>1.18</td>
<td>1.32</td>
<td>1.73</td>
<td>1.54</td>
<td>1.73</td>
<td>2.03</td>
<td>1.54</td>
<td>1.73</td>
<td>2.03</td>
<td>1.54</td>
<td>1.73</td>
<td>2.03</td>
</tr>
</tbody>
</table>
sum of the school and school class variances. However, using only the school-level variance allows us to obtain information about the relative importance of the school versus the school class environment for understanding medication use. In the present investigation, we were mostly interested in measuring the association between bullying and medication use, rather than analyzing the variance of medicine use between schools and between school classes. Therefore, we used multilevel regression analyses only to account for the hierarchical structure of the data, and we did not attempt to explain statistically the school- and school class-level variances.30

Sensitivity analyses were conducted (data not shown), and our results were robust with respect to changes in the definition of the outcome medicine use. Patterns of findings were similar whether they were based on comparisons of the extreme categories of medicine use (yes, several times, versus no) or the categories of yes, once/yes, several times, versus no.

Participants who failed to provide full information on exposure to bullying, prevalence of symptoms, or medicine use were excluded from the analyses (156 boys and 165 girls). There were very few missing data regarding exposure to bullying (0.7%), medicine use (between 1.7% for medicine use for headache and 4.5% for medicine use for nervousness and sleep difficulties), and symptom prevalence (between 1.9% for headache and 4.0% for nervousness), and the final analyses included 2425 boys and 2459 girls.

**RESULTS**

Table 3 shows the gender-specific distributions of bullying, prevalence of medicine use, and prevalence of symptoms and the gender-specific distributions of family social class and age group. For both boys and girls, approximately one half of the pupils had experienced bullying within the current school term; 7.8% of the girls and 8.9% of the boys were victims of bullying every week.

The highest prevalence of medicine use was found for medicine to alleviate headache (34.3% for boys and 45.7% for girls), whereas use of medicine for the other 3 ailments showed much lower prevalence (from 20.4% for medicine use for stomachache among girls to 2.8% for medicine for nervousness among boys). More than boys used medicines for nervousness among boys, whereas use of medicine for headache among girls to 2.8% for medicine use for nervousness among girls to 2.8% for medicine for nervousness among boys. The lowest prevalence of medicine use for headache (2.8% for boys and 5.7% for girls; p = .0781).

For boys, the highest monthly symptom prevalence (52.3% for stomachache), compared with 36.8% for headaches and 5.8% for nervousness, was found for stomachache. For girls, the prevalence of symptoms was approximately the same for all 4 symptoms.
The results were attenuated for medicine for sleep difficulties and nervousness when we adjusted for age group and social class (model 2). Even when we took the higher frequency of symptoms among frequently bullied girls into account (model 3), there was greater frequency of medicine use among bullied girls (ORs between 1.13 [95% CI: 0.8–1.7] for medicine for stomachache and 2.40 [95% CI: 1.5–3.8] for medicine for nervousness).

The MOR measure indicated considerable variance between schools and between school classes, especially for use of medicine for sleep difficulties and medicine for nervousness. In general, the variance was not affected greatly when the individual characteristics of the children were considered. The school class environment seemed to play a more relevant role (i.e., higher MOR) than the school environment.

**DISCUSSION**

This study demonstrated that exposure to bullying not only was associated positively with the prevalence of headache, stomachache, difficulties in getting to sleep, and nervousness; exposure to bullying was also associated positively with the use of medicines for treatment of these symptoms, even after adjustment for the higher prevalence of symptoms among the bullied children. The association between bullying and medicine use was attenuated somewhat with control for symptom prevalence, but it remained significant, except for use of medicine for stomachache. Among girls, use of medicine for stomachache seems to be largely unrelated to bullying. Menstrual pain probably accounts for the largest proportion of the girls’ use of this type of medicine, because we found a strong, graded, positive association between this type of medicine use and age. Among boys, there was a strong negative association between age and use of medicine for stomachache. Although it was insignificant, there was a 45% increase in the odds of using medicine for stomachache among the bullied boys, compared with the nonbullied boys.

The most prevalent type of medicine use in this age group was use of medicine for headache, and our results showed 40% to 70% increased odds of use of this type of medicine among bullied children. The ORs for the effect of bullying on medicine use were higher for the less-prevalent types of medicine use, such as use of medicine for difficulties in getting to sleep, for which bullied children had ORs of 2.40, compared with nonbullied children. However, whereas only 3% to 4% of the children used medicine for difficulties in getting to sleep, we are most concerned with the somewhat low odds of use of medicine for headache among the bullied children (ORs of 1.73 for boys and 1.43 for girls), because this increased risk affects the greatest number of children. More than one third of the boys and almost one half of the girls used medicine for headaches, and our analyses showed

---

**TABLE 3 Basic Information on the Included Variables According to Gender**

<table>
<thead>
<tr>
<th></th>
<th>Proportion, % (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Boys (n = 2581)</td>
</tr>
<tr>
<td>Have experienced bullying this term</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>48.1 (45.7–50.4)</td>
</tr>
<tr>
<td>Once or twice/a few times</td>
<td>43.1 (40.8–45.4)</td>
</tr>
<tr>
<td>At least weekly</td>
<td>8.9 (7.6–10.2)</td>
</tr>
<tr>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Have used medicine in past month for</td>
<td></td>
</tr>
<tr>
<td>Headache</td>
<td>34.3 (32.7–37.1)</td>
</tr>
<tr>
<td>Stomachache</td>
<td>6.2 (5.3–7.7)</td>
</tr>
<tr>
<td>Difficulties in getting to sleep</td>
<td>3.5 (2.8–4.6)</td>
</tr>
<tr>
<td>Nervousness</td>
<td>2.8 (2.2–3.8)</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
</tr>
<tr>
<td>Reported symptom once per month or more</td>
<td></td>
</tr>
<tr>
<td>Headache</td>
<td>48.9 (46.6–51.2)</td>
</tr>
<tr>
<td>Stomachache</td>
<td>30.9 (28.8–33.1)</td>
</tr>
<tr>
<td>Difficulties in getting to sleep</td>
<td>50.8 (48.5–53.1)</td>
</tr>
<tr>
<td>Nervousness</td>
<td>52.3 (50.0–54.6)</td>
</tr>
<tr>
<td>Family social class</td>
<td></td>
</tr>
<tr>
<td>V/H</td>
<td>25.7 (23.7–27.7)</td>
</tr>
<tr>
<td>III/IV</td>
<td>46.3 (44.0–48.6)</td>
</tr>
<tr>
<td>V/VI</td>
<td>15.1 (13.5–16.8)</td>
</tr>
<tr>
<td>Not classifiable/missing</td>
<td>12.9 (11.4–14.6)</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
</tr>
<tr>
<td>Age group</td>
<td></td>
</tr>
<tr>
<td>5th grade (11-y-old students)</td>
<td>33.8</td>
</tr>
<tr>
<td>7th grade (13-y-old students)</td>
<td>36.3</td>
</tr>
<tr>
<td>9th grade (15-y-old students)</td>
<td>29.9</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
</tr>
</tbody>
</table>

* The 95% CI was chosen to take into account the clustering of data in the school-based sample.

(ranging from 55.9% for difficulties in getting to sleep to 60.0% for headache). The prevalence of symptoms was higher among girls than among boys, but gender differences were small for the 2 psychological symptoms (nervousness and difficulties in getting to sleep).

Medicine use for all 4 symptoms was more than twice as high among frequently bullied boys, compared with nonbullied boys (ORs between 2.26 [95% CI: 1.7–3.0] for medicine for headache and 4.43 [95% CI: 3.0–6.5] for medicine for nervousness) (Table 1, model 1). Some of this increased risk was accounted for when analyses were adjusted for age group and social class (model 2). However, even with control for these variables and for the higher symptom prevalence among bullied boys, medicine use was higher among boys who were bullied frequently than among nonbullied boys (model 3: ORs between 1.45 [95% CI: 0.9–2.4] for medicine for stomachache and 2.68 [95% CI: 1.7–4.2] for medicine for sleeping difficulties). Also among girls, we found exposure to bullying to be associated with medicine use, except for medicine use for stomachache, but the estimates for the association were lower than those for boys (ORs between 1.11 [95% CI: 0.8–1.6] for medicine for stomachache and 4.10 [95% CI: 2.7–6.4] for medicine for nervousness) (Table 2, model 1).
that a large number of those children were bullied sometimes or often.

Substantial proportions of adolescents use medicines for common health problems. Internationally, our own studies in 28 countries found prevalence rates for use of medicine for headache ranging between 21.1% (Slovak Republic) and 49.9% (Scotland) among boys and between 28.3% (Greenland) and 65.9% (United States) among girls.1 We found only one other study that addressed the issue of bullying and medicine use. Van Cleave and Davis31 studied bullying and victimization among children with special health care needs, among them a group of children using prescribed medicines. Those authors found the odds of being bullied among those children to be 1.09, compared with children without special health care needs.

It is important to identify factors that influence medicine use in this age group and to understand why bullying is associated with medicine intake. Our study is cross-sectional and does not reveal how bullying may contribute to medicine use, but there are several possible explanations for the association.

First, bully victimization has severe consequences for the health and well-being of the children involved. Studies have shown associations between bullying and a plethora of health-related factors, including physical symptoms, overweight, psychological symptoms, bed-wetting, poor thriving, low self-esteem, social marginalization, and even suicidal ideation.5–10 Although most of those studies were cross-sectional, the causal direction for a number of those associations has been confirmed through longitudinal studies. For instance, longitudinal studies have confirmed that bullying leads to increased risk of physical symptoms, whereas the direction of causality is less clear for some psychological factors. Some studies have shown that vulnerable children are at increased risk of being bullied,13,14,31–33 whereas other studies have found that victimization leads to the onset of emotional problems and that emotional problems do not increase the risk of being bullied.12 Part of the explanation for these differences may be the difference in the ages of the studied populations. Studies of young children (5–9 years of age) found vulnerable children to be at increased risk of bullying, whereas, in their study of adolescents (13–14 years of age), Bond et al12 found bullying to be the cause and emotional problems to be the effect. The differences in results may be a consequence of different mechanisms operating at different ages of the children or may reflect loops of effects between bullying and psychological distress. Our cross-sectional data leave the same question of directionality. We think that the increased pressure on children who are bullied may increase their risk of using any means of coping with the distress and that use of medicine may be one way of trying to relieve the strain.

Second, we know from other studies that children with chronic diseases are at greater risk of being bullied.17,31 It is likely that vulnerable children (for example, children with chronic diseases) use more medicine than other children and that the increased level of bullying among medicine users is partly explained by this mechanism. We were unable to control for chronic diseases in our study.

Third, we observed high MORs across schools and school classes in our multilevel analyses, which indicated that medicine use was sensitive to characteristics of the school and the school class. Behavioral patterns of medicine use may be communicated from one child to another. This mechanism is well known for adolescent health behaviors such as smoking and physical activity, and the same mechanism may explain part of the school and class effects that we found for medicine use for headache, sleeping difficulties, and nervousness. Interestingly, we found almost no class effect in medicine use for stomachache, which may indicate that the individuall-level factors (ie, the strong age and symptom effects, especially among girls) account for the school and class variations for this type of medicine use behavior. Rules at the school level, such as the practices of the school staff regarding whether to distribute medicines to children on request, may influence the level of medicine intake at the school. The large differences in mean levels of bullying among schools and classes may represent another possible explanation for the school- and class-level variances in medicine use,11 because children from schools with a high prevalence of bullying may use more medicine to cope with the social strain of the school environment, compared with children from schools with a more pleasant environment.

Fourth, Andersen et al14 found medicine use to cluster with smoking and drunkenness and suggested that medicine use is part of a behavioral pattern of misuse. In our study, medicine use for headache showed a positive association with age, and it is likely that this type of medicine use was related to a pattern of risk behavior clustering among older children; however, the negative associations between other types of medicine use and age make it unlikely that they are involved in this suggested mechanism. To prevent excess medication use in adolescence, it is important to know when adolescents use medications for psychological distress and when they take them for other purposes, such as recreational use.

Longitudinal research is needed to study the causal mechanisms underlying adolescents’ use of medication. Research in this area is challenged by the developmental changes that occur during the adolescent life period and should address not only the reasons for medication use but also any changes in causes over time.

The study was based on a large, representative study population with a high response rate. Validation studies showed the included variables to be valid for age-equiv-
Furthermore, although the response rate was high, the nonparticipants (11.5%) might have higher prevalence rates of both exposure to bullying and medicine use. If this was the case, then we likely underestimated the associations between exposure to bullying and medicine use.

Medicine use among adolescents is a global public health issue that needs to be addressed through research and public health policies. Bush et al.35 made an important start by developing guidelines for teaching children and adolescents about medicine. These guidelines have been adapted by the International Pharmaceutical Federation,36 and their implementation should be brought into focus.

The results of our study indicate that policymakers, health care professionals, and school staff members should be aware that the adolescent victims of bullying are prone to excess use of medicine. Several intervention studies have demonstrated that pedagogical intervention programs help reduce bullying at school.37–39 Our study gives a second reason to take steps to prevent bullying, that is, to avoid harmful patterns of excess medicine use among adolescents.

ACKNOWLEDGMENTS

This work was supported by the Health Insurance Foundation and the 1991 Pharmacy Foundation (Apotekerfonden af 1991).

We thank Mogens Trab Damsgaard, Institute of Public Health, University of Copenhagen, for his support in discussions on analytical strategy. We also thank Charlotte Ørsted Hougaard for her help with the multilevel analyses presented in Tables 1 and 2. We thank Birgit Pallesen and Karen Steenhard for their help with language and the reviewers for their thorough reviews and helpful comments.

REFERENCES

of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. *J Epidemiol Community Health.* 2006;60:290–297


---

**FASTER MENINGITIS TEST IS APPROVED**

“A test to help doctors rapidly distinguish between viral meningitis and less common but more severe spinal cord and brain infections caused by bacteria received federal approval. The Xpert EV test produced by Cepheid can give results in 2 1/2 hours, far less than the week it can now take, the Food and Drug Administration said. Bacterial meningitis can lead to brain damage, hearing loss and even death if it is not treated. Patients with viral meningitis can recover on their own within two weeks, yet they are often treated, ultimately unnecessarily, with antibiotics as a safeguard against bacterial meningitis, the drug agency said.”


Noted by JFL, MD
Mental Health and Social Competencies of 10- to 12-Year-Old Children Born at 23 to 25 Weeks of Gestation in the 1990s: A Swedish National Prospective Follow-up Study

Aijaz Farooqi, MD, PhD, Bruno Hägglof, MD, PhD, Gunnar Sedin, MD, PhD, Leif Gothefors, MD, PhD, Fredrik Serenius, MD, PhD

Divisions of aPediatrics and bChild and Adolescent Psychiatry, Department of Clinical Sciences, University Hospital, Umeå, Sweden; cSection for Pediatrics, Department of Women’s and Children’s Health, Uppsala University, Uppsala, Sweden

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

ABSTRACT

OBJECTIVE. We investigated a national cohort of extremely immature children with respect to behavioral and emotional problems and social competencies, from the perspectives of parents, teachers, and children themselves.

METHODS. We examined 11-year-old children who were born before 26 completed weeks of gestation in Sweden between 1990 and 1992. All had been evaluated at a corrected age of 36 months. At 11 years of age, 86 of 89 survivors were studied and compared with an equal number of control subjects, matched with respect to age and gender. Behavioral and emotional problems, social competencies, and adaptive functioning at school were evaluated with standardized, well-validated instruments, including parent and teacher report questionnaires and a child self-report, administered by mail.

RESULTS. Compared with control subjects, parents of extremely immature children reported significantly more problems with internalizing behaviors (anxiety/depression, withdrawn, and somatic problems) and attention, thought, and social problems. Teachers reported a similar pattern. Reports from children showed a trend toward increased depression symptoms compared with control subjects. Multivariate analysis of covariance of parent-reported behavioral problems revealed no interactions, but significant main effects emerged for group status (extremely immature versus control), family function, social risk, and presence of a chronic medical condition, with all effect sizes being medium and accounting for 8% to 12% of the variance. Multivariate analysis of covariance of teacher-reported behavioral problems showed significant effects for group status and gender but not for the covariates mentioned above. According to the teachers’ ratings, extremely immature children were less well adjusted to the school environment than were control subjects. However, a majority of extremely immature children (85%) were functioning in mainstream schools without major adjustment problems.

CONCLUSIONS. Despite favorable outcomes for many children born at the limit of viability, these children are at risk for mental health problems, with poorer school results.
In the 1990s, major advances in perinatal intensive care and the development of regionalized perinatal care resulted in dramatic increases in the survival rates for extremely immature (EI) children born before 26 completed weeks of gestation. Follow-up studies of school-age outcomes for infants born in the 1980s with extremely low birth weight (ELBW) (<1000 g) showed that these infants had significantly high prevalence rates of low-severity neuropsychological, behavioral, and school difficulties. In a meta-analysis of case-control studies reported from 1980 to 2001 in which cognitive and behavioral outcomes were examined, it was found that school-aged children who were born very preterm exhibited more internalizing, attention, and externalizing problems, although these difficulties were not encountered in all studies. The vast research on behavioral outcomes of children with very low birth weight (VLBW) (<1500 g) or ELBW has revealed that these children are particularly vulnerable to problems related to inattention, hyperactivity, and social difficulties. There are also reports that preterm children with attention-deficit/hyperactivity disorder (ADHD) have significantly worse performance on cognitive tasks and a higher rate of school failure than control subjects. Furthermore, there is some evidence to suggest that, compared with control subjects, VLBW or ELBW adolescents and VLBW young adults have high prevalence rates of anxiety and depression. Apart from study from Australia and an abstract from the United States, we are not aware of reports of behavioral outcomes at school age for children who were born extremely preterm (<28 weeks of gestation) or with ELBW in the 1990s. It is important to establish the mental health of EI children born in the 1990s, because the dramatic increase in their survival rates in that decade might have increased the number of children at risk for adverse long-term outcomes.

In this part of our ongoing longitudinal investigation, the objectives were to examine behavioral outcomes, social competencies, and adaptive functioning in a national cohort of 11-year-old EI children, in comparison with matched, term, normal, control subjects, born in the 1990s and to investigate whether any observed differences could be explained in part by socioeconomic and environmental factors. We also addressed the question of whether correlates of the behavioral outcomes differed between EI children and control children. On the basis of the available literature, we hypothesized that the EI children would have high rates of ADHD and have social and possibly anxiety/depression problems, that there would be no significant difference in aggressive and delinquent behaviors between EI children and control children, and that the relationship between extreme immaturity and behavioral problems might be partially explained by psychosocial risk factors.

METHODS

Study Groups

The methods of tracing and recruitment of participants and other methodologic details have been described elsewhere and are summarized here. The study participants were survivors of a national cohort of 247 consecutive, live-born, EI (<26 weeks of gestation) infants born during the period from April 1990 through March 1992 in the whole of Sweden. Of those 247 EI infants, 89 (36%) survived the neonatal period, and all were alive at the time of the present study (mean age: 11 years). All of the 89 EI survivors were assessed in their neonatal period and at a corrected age of 36 months, when they were enrolled in the “1000 g” national Swedish cohort. The names and addresses of the EI children and their families, including those that had moved abroad, were traced from the records of the Swedish national tax board, where we also confirmed that the child was alive at the time of the present assessment. A letter was then sent to the pediatrician caring for the EI child, asking whether he or she thought it appropriate for the family to be contacted. Three families that had moved to other countries were traced and approached. Once the families had been located and the pediatricians’ permission obtained, the research nurse wrote to the parents, requesting their written permission to send questionnaires to the children’s schoolteachers, the parents, and the children themselves. Once the EI and control families had given their written agreement to participate in the study, they were contacted by the research nurse, who explained the procedures for completing the questionnaires. With the permission of the parents, questionnaires were sent to the children’s teachers and the children themselves, with a letter with relevant instructions for completing the questionnaires. Children completed the questionnaires at school. If the children had difficulty understanding any of the questions, then the teachers explained the questions. Questionnaires from all respondents were returned to the study coordinator at the University of Umeå. Two reminders were sent to nonrespondents and, when possible, an approach was also made by telephone. Missing data from the returned questionnaires were pursued in the same way.

The control group was recruited at the present assessment (mean age: 11.6 years) from the national birth register, through selection of a child with a normal birth weight born at term at the same hospital, of the same gender and nearest in birth date (±7 days) to the EI child. We identified 3 control participants for every index child. Because we aimed to have 1 control child for every EI child, initially we contacted the first of the control families. If the family did not respond or refused to participate, then we approached the second family and, if necessary, the third. Recruiting the control families was a slow process, and this resulted in a control
group that was an average of 8 months older than the EI children. Eighty-six percent (n = 74) of control families were recruited in the first round of invitation. Ten and 2 control families were recruited in the second and third rounds of invitations, respectively. The control group was approached and examined in the same way as the study population.

**Instruments and Measures**

**Assessment of Mental Health**

For assessment of the parents’ and teachers’ perceptions of the children’s behavior, the parents completed the Child Behavior Checklist (CBCL) for ages 4 to 18 years and the teachers completed the analogous Teacher Report Form (TRF). Both forms include items for scoring particular behavior/emotional problems, plus 2 open-ended problem items. The CBCL and TRF differ somewhat according to the home versus school context in which the child is seen, but the scoring profiles enable both instruments to be scored in terms of a common set of 8 problem behaviors. The list contains 118 items on difficult behaviors, all scored 0 (not true), 1 (somewhat or sometimes true), or 2 (very true or often true). Principal-component analyses reveal 8 sets of behaviors, namely, withdrawn, somatic complaints, anxious or depressed, social problems, thought problems, attention problems, delinquent behavior, and aggressive behavior. Principal-factor analyses of the 8 categories produce 2 broad groupings, namely, internalizing, derived from the sum of the items in the first 3 sets, and externalizing, derived from the last 2 (delinquent behavior and aggressive behavior). The remaining 3 categories (social, thought, and attention problems) represent problems that fit either broad grouping. The CBCL and TRF scales are also constructed in terms of items that measure 6 sets of behavioral syndromes based on the criteria of the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Revised (DSM-IV-R), that is, affective disorder, anxiety, somatic problems, ADHD, conduct disorder, and oppositional-defiant disorder. Respondents were asked to base their answers on the preceding 6 months. To provide a common metric across the different tests, the raw scores for the CBCL and TRF problem subscales and the scores for the CBCL competence scales (activity, social, and school scales) were converted to z scores (SD scores). The z scores for the problem and competence scales of the CBCL were computed relative to Swedish normative values being used in Sweden at the time. Swedish reference normative values for calculation of z scores for TRF problem scales were not available; therefore, we computed z scores for TRF problem scales relative to the mean scores for the control subjects of the same gender. For all behavioral scales, higher scores indicated more problematic behavior. For all TRF and CBCL problem subscales, scores above the 90th percentile for the control subjects of the same gender were classified as being in the abnormal range. The percentile distribution of the total CBCL problem scores for our control group was similar to that for a Swedish reference population. Children completed a self-report with a depression self-rating scale (DSRS). The DSRS is an 18-item self-report questionnaire composed of a psychiatric symptom checklist that measures anxiety and depression. The child is asked to rate his or her own situation during the past month, on a 3-point scale. Scores of 2, 1, and 0 refer to most of the time, sometimes, and never, respectively. For the DSRS, scores above the 90th percentile for the control subjects of the same gender were classified as being in the abnormal range. The items of the DSRS were developed after extensive evaluation of the population and were tested against an operational clinical definition of depressive disorder.

**Measurement of Adaptive Functioning and Social Competencies**

Parents provided responses for the following items from the competence scales of the CBCL: general activities (activity scale), which measure the amount and quality of the child’s participation in sports, hobbies, and organized activities; social activities (social scale), which measure participation in organizations and clubs, number of friends and frequency of contacts with friends, and how well the child gets along with the parents, siblings, and other children; and school performance (school scale), which measures performance in academic subjects on a standard scale, grade retentions, any academic or other problems at school, and special education services received by the child at school. Assessments of academic performance and adaptive functioning at school were based on the TRF described by Achenbach. Teachers provided responses for the 5 items on adaptive functioning that measured the child’s performance in academic subjects and 4 adaptive characteristics, as follows. How hard is she or he working? How appropriately is she or he behaving? How much is she or he learning daily? How happy is the pupil? School difficulties were defined as the child repeating a grade and/or using special educational resources (part-time or full-time).

**Assessment of Demographic Characteristics**

Sociodemographic characteristics, including variables such as the parents’ educational level, the family’s disposable income, and the family structure, were assessed with the Nordic Health and Family Questionnaire. The mother’s education was classified into 3 groups, that is, >12 years, 10 to 12 years, or ≤9 years. The cutoff point for high/low education was between the latter group and the others. The family’s disposable income was classified into 6 groups, and the cutoff point for high/low income was between the 2 highest groups and the others. The family structure was defined as a single-parent
or 2-parent family. Any social risk was defined as single-parent family, mother's educational level of ≤9 years, or low family income. The composite social risk index included maternal education (0 for high school or above and 1 for ≤9 years of schooling), family structure (0 for 2-parent family and 1 for single-parent family), and family income (0 for high income and 1 for low income). The composite scale ranged from 0 for the lowest social risk index to 3 for the highest social risk. Socioeconomic status (SES) was measured on the basis of a Swedish socioeconomic classification system. In its most aggregated form, the classification consists of 6 groups. For the purpose of this study, we categorized SES into 3 main groups, that is, laborers (unskilled and semiskilled workers), nonmanual employees (assistant and intermediate nonmanual employees), and professional employees.

Chronic conditions at 11 years of age were defined as neurosensory impairments (NSIs) and medical or psychiatric illnesses with durations of ≥12 months. Identification and characterization of NSIs in this study population were described previously. Chronic medical conditions were defined as NSIs and medical illnesses excluding psychiatric illnesses. Formal psychometric tests were not conducted but had been performed for all children who were receiving full-time special education (13 EI children and 4 control children). Major disability was defined as moderate or disabling cerebral palsy (CP), severe visual impairment (including unilateral or bilateral blindness), sensorineural disability requiring a hearing aid, or a need for full-time special education in a special school (as a measure of severe mental retardation).

Assessment of Environmental/Family Risk
Family function was measured with a 12-item general functioning scale derived from the Nordic Health and Family Questionnaire. This instrument assesses family functioning in 5 dimensions, that is, affective responsiveness, problem-solving, affective involvement, communication and roles, and behavior control. The items concerning family functioning have 5 alternatives, graded from strongly agree to strongly disagree. Total scores range from 0 to 60, with higher scores indicating disturbance. Maternal mood and psychological health were measured with a 12-item brief inventory devised by a research group in the Division of Psychiatry of Umeå University. On this scale, there are 6 items that measure maternal mood, that is, 3 items measuring positive affect (scores range from 0 to 18) and 3 items measuring negative affect (scores also range from 0 to 18). Low maternal mood was identified when the negative mood score was greater than the positive mood score. The positive mood affect was scored as 0 and the negative affect as 1. The remaining 6 items measure categorical responses to questions on the psychological health of the mother (visiting a psychiatrist or a psychologist, suicidal ideation, use of psychopharmacological drugs, phobias, and fears). Maternal health risk scores included all categorical responses from the maternal mood scale (0 for positive and 1 for negative maternal mood) and the psychological health scale (0 for not true and 1 for true). These scores were added to obtain a composite maternal mental health risk index (ranging from 0 to 12), with higher scores indicating disturbance. The study was approved by the regional ethical review board at Umeå University.

Statistical Analyses
Data were collected on standardized forms and encoded for computerized analysis with the use of Windows SPSS 13.0 (SPSS, Chicago, IL). The assessment data for each EI child were examined before they were combined with the data set from 2 previous main studies for analysis. Descriptive statistics such as frequency distributions, means, and SDs were used. Differences in dichotomous outcomes between the groups were analyzed with χ² tests or Fisher's exact test, when appropriate. Continuous outcome measures were compared with unpaired Student's t tests, to test the differences between the means. Behavioral/emotional problems and competencies are likely to be related to stresses in a child's life, as well as factors such as social risk; therefore, in the analyses of the CBCL and TRF competence and problem scale scores, we controlled for social risk factors, family function, maternal mental health risk, and presence of a chronic medical condition in the child. Multivariate analysis of covariance (MANCOVA) was performed to test the differences between the groups. The independent variables were group status (EI versus control) and gender, and the dependent variables were each of the 6 DSM-IV-R-oriented syndrome scales of the CBCL or analogous scales of the TRF. The following variables were included as covariates/factors in the model: composite social risk index, family function, maternal mental health risk score, and presence of a chronic medical condition. MANCOVAs of CBCL and TRF syndrome scales were performed in 2 separate analyses. Follow-up univariate analysis of covariance (ANCOVA) was conducted for each of the 6 DSM-IV-R-oriented behavioral syndrome subscales of CBCL and TRF, with the same covariates as in the MANCOVAs of CBCL and TRF syndrome scale scores. ANCOVAs were also performed on the 5 TRF adaptive functioning scales (academic performance, daily learning, hard work, classroom behavior, and being happy), the 3 competence scales of CBCL (social, activity, and school scales), and the DSRS, with the same covariates/factors as in the ANCOVAs of the CBCL and TRF syndrome scale scores. The effect size (ES) is given by the partial η² statistic, which describes the proportion of total variability attributable to a factor or covariate (the proportion that, if multiplied by 100, is the percentage of total variability attributable to the
group differences). All significant MANCOVA and follow-up ANCOVA effects were interpreted by using the criteria for ES described by Cohen, with which effects are deemed small, medium, and large if they account for 1% to 5.8%, 5.9% to 13.8%, and >13.8% of variance, respectively.

Multivariate logistic regression analyses were also performed to examine differences in dichotomous behavioral outcomes between the groups. Social risk score, family function, gender, maternal mental health risk score, and presence of a chronic medical condition were entered as covariates. \( P \) values of <.05 were considered significant.

RESULTS

Participants

At the time of this assessment, 1 child had emigrated and was lost to follow-up monitoring and 2 families refused to participate. Therefore, 86 EI children remained for assessment (mean age: 10.9 years; SD: 0.76 years). Of the nonparticipating children, 1 had shown a significant NSI and the other 2 had no disability at 3 years of age. The CBCL and TRF were completed by parents and teachers, respectively, for 83 children in the EI group. For the remaining 3 EI children, the neuropsychological and behavioral assessments could not be performed because of severe impairment. However, other information was available for those children from teacher and parent questionnaires (eg, information on special education needs, special schools, personal assistants, and need for special teachers). Of the 86 questionnaires completed by EI parents, 41 were completed by mothers, 44 by both parents, and 1 by a father. The corresponding numbers for control parents were 53, 32, and 1, respectively. Five EI children were unable to complete self-reports (3 were severely impaired and 2 returned incomplete questionnaires). For all 86 control children, complete CBCL, TRF, and self-report data were obtained.

Demographic Features

Sociodemographic characteristics, infant data, and rates of chronic conditions are shown in Table 1. At 11 years of age, 13 EI children (15%) had NSIs, which included ≥1 of the following conditions: CP for 5, severe visual impairment (including unilateral or bilateral blindness) for 10, and sensorineural disability requiring a hearing aid for 5. In the control group, the corresponding rate was 2% (\( n = 2 \); 1 child had CP, and 1 had severe visual impairment). Of the 86 EI children, 73 (85%) were in mainstream schools and 13 (15%) were receiving full-time special education. The corresponding rates for the control group were 82 (95%) and 4 (5%). The overall prevalence of ≥1 major disability was 21% for the EI children and 6% for the control participants (\( \chi^2 = 7.03; P = .006 \)). There were no statistically significant differences between the EI and control participants regarding family structure, maternal education, maternal mental health risk index, SES, and family function (Table 1).

Mental Health Measures According to Group and Evaluation Source

Psychometric Analyses

A high level of reliability was demonstrated in the analyses of behavioral subscale scores from parents’ reports, teachers’ reports, and children’s self-reports (Cronbach’s \( \alpha = .94, \alpha = .93, \) and \( \alpha = .84, \) respectively). The correlations among the behavior problem subscale scores from the CBCL and TRF ranged from \( r = 0.20 \) to \( r = 0.78 \) and from \( r = 0.18 \) to \( r = 0.75 \), respectively. Most of the subscales of the CBCL or TRF were correlated moderately with one another, with \( r \) values between 0.3 and 0.55. The mean ± SD \( Q \) correlation obtained between the CBCL item scores and the scores obtained for the corresponding item of TRF was 0.26 ± 0.19, indicating an average level of cross-agreement between teachers and parents.

Behavioral Scores According to Parents’ and Teachers’ Reports

Parents and teachers reported significantly higher scores for EI children, compared with control subjects, for internalizing problems (anxious/depressed, withdrawn, or somatic problems) and attention, thought, and social problem scales. The mean \( z \) scores for EI versus control children according to the parent reports were 0.44 vs −0.17 (\( P = .002 \)) for withdrawn behavior, 0.70 vs −0.14 (\( P < .001 \)) for anxious/depressed behavior, 1.17 vs 0.15 (\( P < .001 \)) for attention difficulties, 1.3 vs 0.39 (\( P = .001 \)) for social problems, and 0.75 vs 0.1 (\( P = .002 \)) for thought problems (Fig 1A). The pattern of findings was very similar for the teachers (Fig 1B). On the basis of both parents’ and teachers’ ratings, the mean raw total problem scores were increased significantly for EI children, compared with control children (CBCL score: EI: 23.1; control: 13.72; mean difference: 9.36; \( P = .001 \); TRF score: EI: 25.89; control: 15.05; mean difference: 10.84; \( P = .001 \)). Compared with those of control children, parents and teachers of EI children were more likely to rate the child as scoring in the abnormal range for a number of behaviors (defined as CBCL and TRF problem scale scores above the 90th percentile for control children) (Table 2). The parent-reported rate of abnormal behavior among EI children for the anxious/depressed problem subscale was 27% (adjusted odds ratio [OR]: 2.56; \( P = .036 \)), that for withdrawn problems was 36% (OR: 2.9; \( P = .011 \)), and that for attention problems was 30% (OR: 3.5; \( P = .007 \)). Teachers, like the parents of EI children, reported significantly higher rates of abnormal scores for internalizing (anxious/depressed or withdrawn) and attention problems for EI
children, compared with control children. Teachers also reported significantly higher rates of abnormal scores for EI children for somatic, thought, and social problems (Table 2).

Depression Symptoms According to Children’s Self-Reports (DSRS)
The mean scores in the children’s self-reports (DSRS) were significantly higher for EI children, indicating a trend toward increased depression symptoms, compared with the control group (EI: 7.04; control: 5.45; mean difference: 1.58; 95% confidence interval: 0.29–2.86; P = .017). Five of 18 of the DSRS items were significantly more likely to be reported by EI children, compared with the control participants, with these being “haven’t lots of energy” (P = .003), “horrible dreams” (P = .01), “can’t stick up for myself” (P = .001), “feeling very sad” (P = .04), and “feeling very bored” (P = .02). The proportion of EI children in the abnormal range (defined as a DSRS score above the 90th percentile for the control group)
A and B. Mean $z$ scores for 8 behavior problem scales for EI children, obtained from the parent-reported CBCL\textsuperscript{28} (A) or the teacher-reported TRF\textsuperscript{29} (B). The null lines represent the $z$ score for the Swedish reference population (A) or the $z$ score for the control subjects (B). C. Comparison of problem scales, according to parents’ reports (CBCL), for ELBW or EI children from the 5 population cohorts of 5 countries. The null line represents the $z$ score for the country-specific reference group.\textsuperscript{7} $a$ $P < .005$; $b$ $P < .05$; $c$ statistically significant difference between each ELBW or EI cohort and its country-specific reference group. NS indicates not significant.

FIGURE 1
was not significantly different from the proportion of control children (Table 2). Univariate analyses revealed that prematurity had a small but significant effect on depression symptoms reported by the children, which suggested that the trend toward increased depression symptoms persisted among EI children after adjustment for socioeconomic and other environmental factors (Table 3). None of the other covariates was associated with children’s depression symptoms.

MANCOVA Effects According to Parents’ Reports (CBCL)

The 2 (group status) × 2 (gender) MANCOVA of 6 DSM-IV-R-oriented CBCL syndrome scales revealed significant multivariate main effects for group status (prematurity versus control: Wilks’ λ = 0.88; F_{6,152} = 3.54; P = .003; ES = 12%; social risk: Wilks’ λ = 0.90; F_{6,152} = 2.77; P = .014; ES = 10%; family function: Wilks’ λ = 0.91; F_{6,156} = 2.31; P = .036; ES = 8.2%; presence of a chronic medical condition: Wilks’ λ = 0.92; F_{6,156} = 2.25; P = .042; ES = 7.9%). No interactions emerged, which indicated that prematurity, family function, social risk, and presence of a chronic medical condition were associated independently with the multivariate measure of CBCL syndrome scales.

MANCOVA Effects According to Teachers’ Reports (TRF)

As with the parent reports (CBCL), MANCOVA of DSM-IV-R-oriented analogous TRF syndrome scales revealed a significant multivariate effect for prematurity (Wilks’ λ = 0.88; F_{6,152} = 3.47; P = .003; ES = 13%), which indicated that, across the 6 syndrome scales, there was a significant difference between the EI and control children. There was also a significant multivariate main effect for gender (Wilks’ λ = 0.85; F_{6,152} = 4.46; P < .001; ES = 15%). No interactions emerged, which indicated that prematurity and gender were associated independently with the multivariate measure of TRF syndrome scales.

Univariate Effects According to Parents’ Reports

Group Status (Prematurity Versus Control)

In the follow-up univariate analyses, all of the categorical independent variables and covariates/factors were the same as used in the MANCOVAs of DSM-IV-R-oriented syndrome scales of the CBCL and TRF. The analyses revealed significant effects of prematurity on 3 (affective problems, anxiety, and attention difficulties) of the 6 DSM-IV-R-oriented syndrome scales of the

### Table 2: Frequencies of Abnormal Behaviors Above 90th Percentile Cutoff Values

<table>
<thead>
<tr>
<th></th>
<th>EI (N = 83)</th>
<th>Control (N = 86)</th>
<th>Adjusted OR (95% Confidence Interval)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parents’ reports (CBCL)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxious/depressed</td>
<td>22 (27)</td>
<td>9 (10)</td>
<td>2.56 (1.06–6.18)</td>
<td>.036</td>
</tr>
<tr>
<td>Withdrawn</td>
<td>30 (36)</td>
<td>12 (14)</td>
<td>2.9 (1.27–6.63)</td>
<td>.011</td>
</tr>
<tr>
<td>Somatic complaints</td>
<td>11 (13)</td>
<td>7 (8)</td>
<td>1.26 (0.42–3.72)</td>
<td>.68</td>
</tr>
<tr>
<td>Social problems</td>
<td>21 (25)</td>
<td>10 (12)</td>
<td>1.92 (0.79–4.63)</td>
<td>.14</td>
</tr>
<tr>
<td>Thought problems</td>
<td>16 (20)</td>
<td>9 (10)</td>
<td>1.78 (0.71–4.5)</td>
<td>.22</td>
</tr>
<tr>
<td>Attention problems</td>
<td>25 (30)</td>
<td>8 (9)</td>
<td>3.46 (1.40–8.54)</td>
<td>.007</td>
</tr>
<tr>
<td>Aggressive behavior</td>
<td>11 (13)</td>
<td>10 (12)</td>
<td>0.99 (0.36–2.73)</td>
<td>.98</td>
</tr>
<tr>
<td>Delinquent behavior</td>
<td>9 (11)</td>
<td>9 (10)</td>
<td>0.87 (0.31–2.49)</td>
<td>.80</td>
</tr>
<tr>
<td>Internalizing</td>
<td>27 (33)</td>
<td>9 (10)</td>
<td>3.35 (1.38–8.11)</td>
<td>.007</td>
</tr>
<tr>
<td>Externalizing</td>
<td>8 (10)</td>
<td>7 (8)</td>
<td>0.76 (0.22–2.61)</td>
<td>.66</td>
</tr>
<tr>
<td>Total problems</td>
<td>24 (29)</td>
<td>9 (10)</td>
<td>2.86 (1.17–7.0)</td>
<td>.021</td>
</tr>
<tr>
<td><strong>Teachers’ reports (TRF)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxious/depressed</td>
<td>19 (23)</td>
<td>8 (9)</td>
<td>3.54 (1.39–9.03)</td>
<td>.008</td>
</tr>
<tr>
<td>Withdrawn</td>
<td>19 (23)</td>
<td>8 (9)</td>
<td>3.15 (1.25–8.0)</td>
<td>.15</td>
</tr>
<tr>
<td>Somatic complaints</td>
<td>17 (21)</td>
<td>6 (7)</td>
<td>3.94 (1.37–11.32)</td>
<td>.11</td>
</tr>
<tr>
<td>Social problems</td>
<td>17 (21)</td>
<td>7 (8)</td>
<td>2.86 (1.08–7.58)</td>
<td>.035</td>
</tr>
<tr>
<td>Thought problems</td>
<td>25 (30)</td>
<td>7 (8)</td>
<td>5.04 (1.87–13.61)</td>
<td>.001</td>
</tr>
<tr>
<td>Attention problems</td>
<td>20 (24)</td>
<td>6 (7)</td>
<td>3.43 (1.26–9.35)</td>
<td>.016</td>
</tr>
<tr>
<td>Aggressive behavior</td>
<td>17 (21)</td>
<td>11 (13)</td>
<td>1.33 (0.53–3.33)</td>
<td>.53</td>
</tr>
<tr>
<td>Delinquent behavior</td>
<td>19 (23)</td>
<td>9 (10)</td>
<td>2.20 (0.89–5.45)</td>
<td>.08</td>
</tr>
<tr>
<td>Internalizing</td>
<td>21 (25)</td>
<td>8 (9)</td>
<td>3.51 (1.41–8.78)</td>
<td>.007</td>
</tr>
<tr>
<td>Externalizing</td>
<td>15 (18)</td>
<td>9 (10)</td>
<td>1.76 (0.65–4.76)</td>
<td>.27</td>
</tr>
<tr>
<td>Total problems</td>
<td>20 (24)</td>
<td>8 (9)</td>
<td>3.1 (1.19–8.07)</td>
<td>.021</td>
</tr>
<tr>
<td><strong>Children’s self-reports (DSRS)</strong></td>
<td>10 (12)</td>
<td>8 (9)</td>
<td>1.27 (0.46–3.54)</td>
<td>.64</td>
</tr>
</tbody>
</table>

Data were adapted from the problem scales of the CBCL and TRF and from the children’s self-reports (DSRS). Cutoff values for all scales were >90th percentile for control children of the same gender.

*The ORs were derived from logistic regression analyses adjusted for gender, social risk, family function, maternal mental health risk score, and presence of a chronic medical condition.

Eighty-one EI children and 86 control children had complete DSRS data.
CBCL (Table 3), which suggested that the significant behavior problems persisted in these domains among the EI children after controlling for psychosocial risk factors, gender, and the presence of a chronic medical condition. The analyses also demonstrated that, in addition to prematurity, a number of other variables were associated significantly with behavior problems, as described below.

**Family Function**

Effects of family function emerged for the following domains: anxiety problems ($F_{1,161} = 8.87; P = .003; ES = 5.2\%$), somatic problems ($F_{1,161} = 4.61; P = .03; ES = 2.8\%$), and attention problems ($F_{1,161} = 4.03; P = .046; ES = 2.4\%$).

**Gender**

Gender effects emerged for ADHD problems ($F_{1,161} = 6.2; P = .014; ES = 3.7\%$), indicating that boys had higher ADHD scores than girls. However, there was no interaction between status and gender ($F_{1,161} = 0.541; P = \text{not significant}$), which suggested that EI boys and girls differed from their respective control subjects to similar extents.

**Chronic Medical Condition**

The presence of a chronic medical condition was associated with affective problems ($F_{1,161} = 9.82; P = .002; ES = 5.8\%$) and anxiety problems ($F_{1,161} = 6.32; P = .013; ES = 3.8\%$).

**Univariate Effects According to Teachers' Reports**

ANCOVAs of 6 DSM-IV-R-oriented syndrome scales of the TRF revealed significant effects of prematurity on 4 disorders, namely, affective, anxiety, somatic, and attention difficulties (Table 3). As with the CBCL, a significant effect of gender on ADHD problems emerged ($F_{1,158} = 20.23; P < .001; ES = 11.4\%$) and there was no interaction of group status with gender ($F_{1,158} = 1.835; P = \text{not significant}$). No other covariates were associated with any of the DSM-IV-R-oriented syndrome scales of the TRF. In a series of logistic regression analyses based on parent- and teacher-reported abnormal behavioral outcomes (CBCL and TRF), we found that adjustment for demographic variables (social risk, family function, maternal mental health, presence of a chronic medical condition, and gender) made no material difference to the results, which remained significantly high in the EI cohort (Table 2).

**Adaptive Functioning and Social Competencies**

Comparisons between the EI and control cohorts were made with respect to a number of items that reflect adaptive functioning and social competencies. On the TRF scale that directly assesses academic performance in a mainstream school, EI children had significantly

---

**TABLE 3** Parents' Reports, Teachers' Reports, and Children's Self-Reports on Behavioral Problems of EI Children and Control Children at 11 Years of Age

<table>
<thead>
<tr>
<th></th>
<th>Mean ± SD</th>
<th>F Ratio</th>
<th>P</th>
<th>ESa</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EI (N = 83)</td>
<td>Control (N = 86)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parents' reports (CBCL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total problem scores</td>
<td>23.1 ± 19.94</td>
<td>13.72 ± 16.1</td>
<td>5.14</td>
<td>.025</td>
</tr>
<tr>
<td>Internalizing scores</td>
<td>7.22 ± 7.19</td>
<td>3.48 ± 3.75</td>
<td>10.53</td>
<td>.001</td>
</tr>
<tr>
<td>Externalizing scores</td>
<td>6.15 ± 7.03</td>
<td>5.20 ± 7.07</td>
<td>0.007</td>
<td>.93</td>
</tr>
<tr>
<td>Affective problems</td>
<td>1.91 ± 2.51</td>
<td>0.78 ± 1.28</td>
<td>7.50</td>
<td>.007</td>
</tr>
<tr>
<td>Anxiety problems</td>
<td>1.80 ± 2.26</td>
<td>0.67 ± 1.21</td>
<td>10.92</td>
<td>.001</td>
</tr>
<tr>
<td>Somatic problems</td>
<td>1.49 ± 1.83</td>
<td>0.86 ± 1.27</td>
<td>2.91</td>
<td>.09</td>
</tr>
<tr>
<td>ADHD problems</td>
<td>2.66 ± 2.56</td>
<td>1.41 ± 2.26</td>
<td>6.83</td>
<td>.01</td>
</tr>
<tr>
<td>Conduct problems</td>
<td>1.17 ± 2.20</td>
<td>1.0 ± 2.06</td>
<td>0.04</td>
<td>.83</td>
</tr>
<tr>
<td>Oppositional-defiant problems</td>
<td>2.15 ± 2.24</td>
<td>1.87 ± 2.29</td>
<td>0.001</td>
<td>.97</td>
</tr>
</tbody>
</table>

**Teachers' reports (TRF)**

<table>
<thead>
<tr>
<th></th>
<th>Mean ± SD</th>
<th>F Ratio</th>
<th>P</th>
<th>ESa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total problem scores</td>
<td>25.89 ± 22.58</td>
<td>15.05 ± 20.5</td>
<td>8.97</td>
<td>.003</td>
</tr>
<tr>
<td>Internalizing scores</td>
<td>9.22 ± 8.24</td>
<td>7.08 ± 8.31</td>
<td>10.85</td>
<td>.001</td>
</tr>
<tr>
<td>Externalizing scores</td>
<td>5.78 ± 8.26</td>
<td>4.52 ± 7.38</td>
<td>0.828</td>
<td>.36</td>
</tr>
<tr>
<td>Affective problems</td>
<td>1.62 ± 2.37</td>
<td>0.79 ± 1.88</td>
<td>5.87</td>
<td>.017</td>
</tr>
<tr>
<td>Anxiety problems</td>
<td>1.87 ± 2.11</td>
<td>0.88 ± 1.75</td>
<td>11.57</td>
<td>.001</td>
</tr>
<tr>
<td>Somatic problems</td>
<td>1.01 ± 1.46</td>
<td>0.48 ± 1.23</td>
<td>5.08</td>
<td>.026</td>
</tr>
<tr>
<td>ADHD problems</td>
<td>4.59 ± 5.56</td>
<td>2.82 ± 4.80</td>
<td>4.49</td>
<td>.036</td>
</tr>
<tr>
<td>Conduct problems</td>
<td>1.04 ± 2.21</td>
<td>0.71 ± 1.90</td>
<td>1.14</td>
<td>.28</td>
</tr>
<tr>
<td>Oppositional-defiant problems</td>
<td>1.32 ± 1.87</td>
<td>1.27 ± 1.79</td>
<td>0.006</td>
<td>.94</td>
</tr>
</tbody>
</table>

**Children's self-reports (DSRS)**

<table>
<thead>
<tr>
<th></th>
<th>Mean ± SD</th>
<th>F Ratio</th>
<th>P</th>
<th>ESa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total problem scores</td>
<td>7.04 ± 4.51</td>
<td>5.45 ± 4.01</td>
<td>4.31</td>
<td>.04</td>
</tr>
</tbody>
</table>

Means were adjusted for the effects of child’s gender, social risk, maternal mental health risk, family function, and presence of a chronic medical condition. Higher scores indicate poorer functioning.

ES is given by the partial $\eta^2$ statistic (see “Methods”).

Data were obtained from the children’s self-reports (DSRS)$^{32}$; 81 EI children and 86 control children had complete DSRS data.
poorer scores than control participants ($P < .001$) (Table 4). ANCOVA revealed that the group status (prematurity) represented a large effect, accounting for 22% of the variance, which indicated that prematurity was associated strongly with poor academic performance, as indicated by teacher ratings. None of the other covariates was associated with the academic performance of either EI or control children. Furthermore, the EI children had significantly poorer scores than the control children with respect to TRF ratings of how much the child was learning ($P < .001$) and total adaptive function, which was computed by summing ratings for 4 adaptive characteristics ($P < .001$) (Table 4). There were no statistically significant interactions between gender and group regarding these variables. On the CBCL school and social scale, the EI children obtained poorer scores than the control participants (Table 4). Social risk was associated with the CBCL activity scale ($F_{1,162} = 11.29; P = .001$; ES = 6.5%). Although social risk was also associated with the CBCL social scale, it represented only a small effect (4% of the variance), compared with the medium effect (8%) observed for group status (prematurity versus control). Compared with control children, significantly larger proportions of EI children had failed a grade (15% vs 5%; $\chi^2 = 6.27; P < .05$) or had school difficulties (59% vs 12%; $\chi^2 = 41.23; P < .001$) (Table 5). Thirteen EI children, compared with 4 children in the control group ($P = .023$), were receiving full-time special education in a special school or special class. A significantly smaller proportion of EI children, compared with control children, participated in any sports at school or at home (76% vs 92%; $\chi^2 = 7.38; P = .004$) and significantly more EI children were rated below average in sports (29% vs 5%; $\chi^2 = 17.53; P < .001$). Of the 21 EI children who did not participate in sports, 11 children had NSIs (CP for 4, visual difficulties for 6, and deafness for 1), 7 tired easily because of poor motor skills and poor coordination, and 3 had ADHD problems. However, no differences were found in the proportions of children who participated in activities and games other than sports. Furthermore, the proportions of children with impaired family relationships or relationships with peers and those who had few friends did not differ between the 2 groups (Table 5).

**DISCUSSION**

In this population-based study, on the basis of parents’ and teachers’ ratings, the EI children had significantly higher problem scores for internalizing behaviors (anxiety/depression, withdrawn, and somatic problems) and attention, social, and thought problems. The children’s self-reports also pointed to an increased trend toward depression symptoms in the EI cohort. No differences in externalizing problems were indicated by either the parents’ or teachers’ reports. The present study incorporated several methodologic features that strengthened the validity of our findings. The notable strengths included the national composition of the study, with a very high follow-up rate (97%), the prospective follow-up monitoring, and the inclusion of the matched group of children who were born at the same hospital and nearest in time to the EI children. The use of classroom control groups from mainstream schools to represent a relatively healthy group might have overestimated the difference for preterm children, especially in the frequency of cognitive, behavioral, and school problems. Additional strengths of our study were that we used well-validated instruments to measure a wide range of behavioral problems and impairments in adaptive functioning. CBCL and TRF have been used for both clinical and research purposes for many years, across different cultures. We obtained reports on children’s

**TABLE 4 Adjusted Mean Scores for Adaptive Functioning in School From Teachers’ Reports and Competence Scores From Parents’ Reports**

<table>
<thead>
<tr>
<th>Items</th>
<th>EI</th>
<th>Control</th>
<th>F Ratio</th>
<th>P</th>
<th>ESa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptive functioning scales (TRF) scoreb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic performance</td>
<td>73</td>
<td>2.73 ± 0.69</td>
<td>82</td>
<td>3.46 ± 0.63</td>
<td>41.46</td>
</tr>
<tr>
<td>Works hard</td>
<td>73</td>
<td>3.97 ± 1.49</td>
<td>82</td>
<td>4.80 ± 1.24</td>
<td>16.54</td>
</tr>
<tr>
<td>Behaves appropriately</td>
<td>73</td>
<td>3.88 ± 1.18</td>
<td>82</td>
<td>4.48 ± 1.16</td>
<td>12.18</td>
</tr>
<tr>
<td>Learning</td>
<td>73</td>
<td>3.63 ± 1.25</td>
<td>82</td>
<td>4.70 ± 1.06</td>
<td>35.94</td>
</tr>
<tr>
<td>Happy</td>
<td>73</td>
<td>4.16 ± 1.32</td>
<td>82</td>
<td>4.60 ± 1.14</td>
<td>3.47</td>
</tr>
<tr>
<td>Total adaptive scorec</td>
<td>73</td>
<td>15.64 ± 4.19</td>
<td>82</td>
<td>18.57 ± 3.82</td>
<td>22.49</td>
</tr>
<tr>
<td>Competence scale (CBCL), z score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Activities scale</td>
<td>83</td>
<td>−0.28 ± 1.17</td>
<td>86</td>
<td>−0.01 ± 0.99</td>
<td>2.17</td>
</tr>
<tr>
<td>Social scale</td>
<td>83</td>
<td>−0.62 ± 1.23</td>
<td>86</td>
<td>−0.003 ± 0.98</td>
<td>9.45</td>
</tr>
<tr>
<td>School scale</td>
<td>83</td>
<td>−1.17 ± 1.27</td>
<td>86</td>
<td>−0.067 ± 1.07</td>
<td>31.16</td>
</tr>
</tbody>
</table>

Means were adjusted for the effects of child’s gender, social risk, family function, maternal mental health risk, and presence of a chronic medical condition. Data were adapted from the adaptive functioning scale of the TRFb and the competence scales of the CBCLc.

a ES is given by the partial η2 statistic (see “Methods” for description).

b Data on adaptive functioning at school were analyzed for children who were in mainstream schools (EI: n = 73; control: n = 82).

c The total adaptive score was derived from the sum of 4 adaptive characteristics from the TRF functioning scale,29 as follows. How hard is she or he working? How appropriately is she or he behaving? How much is she or he learning daily? How happy is the pupil?
school functioning, behavior, and social competencies from both teachers and parents. For children at risk of developmental and educational problems, it may be especially crucial to obtain ratings from different adults in different settings, namely, school and home. Each rater can provide important and different information on a child’s behavior, because perceptions and interpretations of behavior vary and may be valid reflections of the child’s behavior, because perceptions and interpretations of behavior vary and may be valid reflections of the child’s behavior in different contexts. One of the main weaknesses of the study was the lack of in-depth psychiatric interviews, which would have allowed for categorization of children according to DSM-IV-R diagnoses of psychopathological conditions. Furthermore, the behavioral data were not obtained through direct observations of the children. Addition of observational assessments to our multiple-informant approach would have strengthened the validity and reliability of the behavioral data for all survivors from 4 large prospective studies of preterm children were obtained with the same instrument, namely, the CBCL. It was found that, despite cross-cultural differences, some behavioral problems and characteristics were strikingly similar for all cohorts. The mean scores for social, thought, and attention difficulty scales were 0.5 to 1.2 SDs higher than country-specific normative values or control values. In none of the cohorts were there any significant differences in the aggressive or delinquent behaviors between the ELBW children and the control children or a country-specific reference group. These findings were strikingly similar to our results; in our EI cohort, the mean z scores for social, thought, and attention problem scales, as rated by parents, were 0.75 to 1.3 SDs higher than those for the control group (Fig 1C). Similar ratings were

### TABLE 5 Parents’ Reports of Adaptive Functioning and Social Competencies According to Group

<table>
<thead>
<tr>
<th></th>
<th>EI</th>
<th>Control</th>
<th>P for Difference Between EI and Control Childrena</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participation in ≥1 sport</td>
<td>65/86 (76)</td>
<td>80/86 (92)</td>
<td>.004</td>
</tr>
<tr>
<td>Competence in sportsb</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below average</td>
<td>19/65 (29)</td>
<td>4/79 (5)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Average</td>
<td>38/65 (59)</td>
<td>55/79 (69)</td>
<td></td>
</tr>
<tr>
<td>Above average</td>
<td>8/65 (12)</td>
<td>21/79 (26)</td>
<td></td>
</tr>
<tr>
<td>Participation in hobbies, activities, or games other than sports</td>
<td>73/86 (85)</td>
<td>77/86 (89)</td>
<td>.51</td>
</tr>
<tr>
<td>Participation in jobs or chores (≥1)</td>
<td>59/86 (69)</td>
<td>67/86 (77)</td>
<td>.23</td>
</tr>
<tr>
<td>Belonging to organizations, clubs, or teams (≥1)</td>
<td>53/86 (62)</td>
<td>70/86 (81)</td>
<td>.007</td>
</tr>
<tr>
<td>Few friends</td>
<td>11/86 (13)</td>
<td>4/86 (5)</td>
<td>.06</td>
</tr>
<tr>
<td>Impaired family relationship</td>
<td>10/84 (12)</td>
<td>5/86 (6)</td>
<td>.18</td>
</tr>
<tr>
<td>Impaired relationship with peers</td>
<td>11/83 (13)</td>
<td>4/86 (5)</td>
<td>.06</td>
</tr>
<tr>
<td>Cannot work or play alone</td>
<td>12/83 (14)</td>
<td>7/86 (8)</td>
<td>.22</td>
</tr>
<tr>
<td>School performance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Special class or special schoolb</td>
<td>13/86 (15)</td>
<td>4/86 (5)</td>
<td>.023</td>
</tr>
<tr>
<td>Grade repetition</td>
<td>13/83 (15)</td>
<td>4/86 (5)</td>
<td>.027</td>
</tr>
<tr>
<td>School difficultiesd</td>
<td>51/86 (59)</td>
<td>10/86 (12)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Data were adapted from CBCL competence scores. P values were calculated with Fisher’s exact test. Ratings of competence in sports were for only those who participated in sports. Attending a special school or training school for the physically disabled and severely mentally retarded or receiving full-time special education attached to the mainstream school. School difficulties were defined as repetition of a grade and/or use of special educational resources (part-time or full-time).

Studies of prematurity-related behavioral outcomes during adolescence for children born before the 1990s disclosed a wide array of emotional problems and behavioral disturbances. ADHD was linked to prematurity. In a meta-analysis, 227 studies reporting behavioral and cognitive data that were published between 1980 and 2001 were reviewed. Among 16 of those studies, presenting behavioral data that were considered worthy of review and involving children born before 1990s, 69% revealed significant increases in internalizing problems, 75% significant increases in externalizing problems, and 67% higher rates of attention problems. In the only population-based study on neurobehavioral outcomes of 8-year-old ELBW children who were born in the 1990s, the children were found to have attention difficulties, internalizing problems, and fewer adaptive skills, compared with the normal birth weight cohort. In our study, both the parents and teachers reported increased internalizing and attention difficulties and significantly lower scores for adaptive functioning for EI children, compared with control children.

In an international study of 408 ELBW children, 8 to 10 years of age, from 4 countries, behavioral problems were compared with a cross-cultural perspective; behavioral data for all survivors from 4 large prospective studies of preterm children were obtained with the same instrument, namely, the CBCL. It was found that, despite cross-cultural differences, some behavioral problems and characteristics were strikingly similar for all cohorts. The mean scores for social, thought, and attention difficulty scales were 0.5 to 1.2 SDs higher than country-specific normative values or control values. In none of the cohorts were there any significant differences in the aggressive or delinquent behaviors between the ELBW children and the control children or a country-specific reference group. These findings were strikingly similar to our results; in our EI cohort, the mean z scores for social, thought, and attention problem scales, as rated by parents, were 0.75 to 1.3 SDs higher than those for the control group (Fig 1C). Similar ratings were
reported by teachers (TRF) for our EI children. Furthermore, our EI children did not display externalizing (aggressive or delinquent) behavior problems, as indicated by parent or teacher report. This finding of an increase in attention/hyperactivity problems but not in conduct disorders or aggressive behaviors is in line with the findings from a number of other studies on behavioral outcomes for ELBW or VLBW children investigated at 8 to 12 years of age,7,9,12,14 in adolescence,19,41 or at a young adult age.22,46 It is particularly surprising that these children, despite having many risk factors for externalizing and risk-taking behavior (ADHD problems, social problems, and learning difficulties), do not have high rates of conduct disorders, according to teachers or parents. Furthermore, these results are not in agreement with those of follow-up studies of children with ADHD problems in the general pediatric population, which showed high rates of continuing psychopathological conditions and comorbidity of conduct and oppositional-defiant disorders.47 It has been postulated that the low rates of risk-taking may be attributable to increased parental monitoring and protection.46,48 An increased risk of anxious/depressed and withdrawn behaviors, as seen in our EI children, might lead to behavioral inhibition and to decreased antisocial behaviors.49 There is some evidence that the ADHD symptoms reported among VLBW children are more of the inattention type than of the hyperactivity type12 and are less associated with comorbidity and conduct disorders.9,12,14 This might also explain why ADHD symptoms in preterm children do not seem to have negative implications for the persistence of major sequelae into adulthood.50 However, it is evident that the ADHD profile seen in our EI cohort and in many other VLBW or ELBW populations does not represent a clinical diagnosis of ADHD. Others have suggested that the interaction between perinatal and social risk factors may be a critical factor.51 More studies are needed to elucidate how child outcomes are influenced by the interaction between parenting and perinatal complications. However, the mental health outcomes in adulthood for extremely preterm children born in the modern era of perinatal care remain to be seen.

The depression rating scale includes items such as “life not worth living,” “feels sad,” “lonely,” “does not sleep well,” “feels bored,” and “cries a lot.”32 Although they do not measure exactly the same construct, the anxious/depressed syndromes of the CBCL and TRF scales comprise similar questions.28,29 Our findings showed increases in anxious/depressed and withdrawn symptoms for EI children according to parents’ and teachers’ reports and a significant increase in the trend toward depression symptoms according to children’s self-reports. Increased rates of depression,13,14,19,22,45 anxiety,13,42 and overall internalizing behaviors, including shy and withdrawn behaviors,44,45 have been reported for VLBW or ELBW children. A direct link between prematurity and isolated/withdrawn behaviors at school has been reported.45,52 According to the mediational model postulated by Nadeau et al,44 it seems that functional limitations that are consequent to the extreme prematurity may explain anxious/withdrawn and isolated behaviors. Middle school age is considered to be full of competitive tasks in a teenager’s life; in particular, physical activities such as team and individual sports are highly valued. In our cohort, more than one fourth of the EI children had poor motor skills, one fourth were restricted in their activities because of ≥1 handicap, and almost one half had moderate/severe learning difficulties. These rates remained significantly high when children without NSIs were analyzed.22 We, like others,44 think that these difficulties might have imposed on our EI children functional limitations in social situations, leading to more-cautious (even anxious and withdrawn) behaviors during interactions with peers. However, in-depth studies are needed to determine the extent to which these limitations create handicaps for these children that lead them to withdraw from their social groups. Our findings are in conformity with those of others who reported that children with chronic medical illnesses were at increased risk for emotional problems, particularly depression/anxiety and peer conflict/social withdrawal.53–55 Furthermore, in our study, as found by others,13,16,19 differences in behavioral problems persisted whether children with NSIs were excluded.

The clinical interpretation of the increased scores for thought problems according to parent and teacher reports in our EI cohort is not well understood. Similar results were obtained in a multinational study of 8- to 10-year-old ELBW children7 and by Hack et al,22 who found higher rates of thought problems for both VLBW men and women, compared with control subjects, in a behavioral study of young adults at the age of 20 years. The CBCL and TRF thought problem scales are relatively short subscales and reflect a heterogeneous group of DSM-IV-R disorders.56 They include items such as “can’t get mind off,” “repeats acts,” “sees things,” “hears things,” “strange behavior,” and “strange ideas,” with the TRF having additional items such as “harms self” and “fears.” An association between schizophrenia and low birth weight has been suggested from a few epidemiological studies, but it is not clear whether this pertains to intrauterine growth failure, prematurity, or both.57–59 There is some evidence suggesting an association between autistic spectrum disorder and prematurity.60 Others have reported an increased incidence of autism in very preterm children with an unfavorable visual status attributable to severe retinopathy of prematurity.61,62 The implications of these behavioral difficulties reported by both parents and teachers for our EI children are not clear. It is likely that these behavioral tendencies will decrease with age as the children grow, as do other
problems such as hyperactivity\textsuperscript{15} and externalizing behavior.\textsuperscript{63}

Sociodemographic and environmental factors also significantly influence the long-term outcomes and quality of life for children born preterm.\textsuperscript{13,19} Saigal et al\textsuperscript{19} found that birth weight, family function, maternal mood, gender, and SES predicted the behavioral adjustment of immature adolescents. Levy-Shiff et al\textsuperscript{13} reported that birth weight and psychosocial/environmental variables were significant predictors of later emotional adjustment. McCormick et al\textsuperscript{17} assessed the contributions of birth weight status and family environment to behavior problems at school. They pointed out the greater importance of environmental factors, compared with birth status, and proposed interventions aimed at reducing the environmental risks. Klebanov et al\textsuperscript{64} found that, beside birth weight status, maternal mental depression and home environment were associated significantly with social competence scores. In our study, boys showed a significantly increased risk for attention problems. Family function, social risk, and presence of a chronic medical condition predicted behavioral adjustment in our study population. Although we did not measure exactly the same constructs, our findings were similar to those of Levy-Shiff et al,\textsuperscript{13} Saigal et al,\textsuperscript{19} and Klebanov et al\textsuperscript{64} with respect to the importance of family environment and social risk in predicting the behavioral adjustment of preterm children. Furthermore, in accordance with previous studies,\textsuperscript{12,19} we did not find any interactions between group status (EI and control groups) and other covariates, which indicated that the correlates of behavioral outcomes did not differ between our EI children and the control children.

In terms of school performance, as reported by parents and teachers, more than one half of the children in our EI cohort were experiencing school problems. However, 85\% of the EI children were attending mainstream schools and the majority of them were not having major adjustment difficulties. Despite there being fewer adaptive skills in our EI cohort, these children were not different from the control children with respect to being happy and being positively adjusted in their daily lives. Although the presence of a chronic medical condition was associated with poor school performance, as reported by parents, it represented only a small effect (3\%), compared with the large effect (16\%) of group status (ie, of being born extremely preterm). No other environmental variables were associated with school problems, as evident from parents and teachers. ADHD problems were a significant predictor of school difficulties, as reported by parents and teachers (data not shown). Results pertaining to executive functions/ADHD problems and school achievement will be the subject of a separate report. Our findings are in agreement with previous reports that showed higher rates of school failure and of grade retentions for ELBW teens, compared with control subjects.\textsuperscript{16,19,65–68} Furthermore, our results are similar to those of a number of other reports that demonstrated significant associations between ADHD and poor academic performance, high rates of school failures (including lower grades at school), and receiving full-time special education.\textsuperscript{11,19–21,66–68} In terms of adaptive functioning, as judged from parent reports, our EI children were seen by their parents as being less well-adjusted in extracurricular activities. These findings could not be explained in terms of more-disabled children who were unable to participate, because these differences persisted when children with NSIs were excluded from the analysis. Others have reported a similar finding of poorer social adjustment, involving mainly social functioning as judged by social attainments and global evaluations of competence, especially in activities such as sports.\textsuperscript{19,65} However, more than two thirds (76\%) of our EI children were participating in coached sports, and no difference was seen between the groups in having few friends or in teacher and peer relationships.

Various plausible mechanisms might explain the association between extreme immaturity and behavioral problems. MRI studies of the brain in immature children have shown a relationship between an abnormal brain structure and certain behavioral difficulties.\textsuperscript{69,70} Perinatal and neonatal complications that lead to central nervous system insults have been shown to be related to ADHD problems.\textsuperscript{10,69} Nagy et al\textsuperscript{71} reported brain MRI findings for 27 preterm children with ADHD problems at 11 years of age: the children had not suffered from intraventricular hemorrhage or periventricular leukomalacia in the neonatal period. It was found that the preterm cohort had significant white matter disturbances and reduced brain volumes, compared with control subjects. There is good evidence that disruptions in cortical development (corticogenesis) and brain connectivity are related inversely to birth weight and gestational age at birth, particularly for very preterm infants, even in the absence of a concomitant biomedical risk.\textsuperscript{22,23} Despite increased knowledge regarding these processes, the functional significance of brain abnormalities is not clear. Social isolation, that is, difficulty in maintaining social relationships and lack of ability to make friends, has been reported to occur more frequently in children who are hyperactive or have attention-deficit disorder.\textsuperscript{74} Life-threatening events in the neonatal period might result in increased parental monitoring and overprotective parental attitudes, leading to deviations in perceptions of the child’s behavior and to inadequate socioemotional adjustment in the child.\textsuperscript{75}

**CONCLUSIONS**

Our results suggested that the EI children had significantly poorer mental health and emotional well-being than did the control participants, including internalizing, attention, social, and thought problems. It is not clear
what these findings imply for late adolescent and young adult life, although one might reasonably speculate that, in the absence of externalizing (antisocial behavior) problems, the prognosis is relatively good. However, this would not decrease the risk for internalizing problems. Our study also showed that a majority of our EL children were functioning in mainstream schools and had no major adjustment problems. Like others,\textsuperscript{19} we think that knowledge regarding the course of mental health problems in childhood and beyond is crucial for identifying the need for intervention and prevention strategies. When there is evidence to suggest attention problems and other behavioral problems, early identification and preventive interventions might help families to manage the problems at an early stage.\textsuperscript{76–78} Our findings are in conformity with reports on a significant relationship between behavioral problems and psychosocial risk factors.\textsuperscript{13,17,19,44,64} Although biological immaturity is associated with a substantial number of behavioral problems, it seems that improvement of the modifiable environmental and socioeconomic risk factors can improve the outcomes for these children. These findings also suggest that current preterm follow-up programs might benefit from the addition of psychological and family services to traditional neurodevelopmental assessment programs, especially in the neonatal period and first years of life.

ACKNOWLEDGMENTS
This study was supported by the Oskarfonden Foundation and the Sven-Jerrings Fond Foundation.

We are indebted to the study children and their families and schoolteachers for their cooperation. We especially thank Dr Hans Stenlund in Umeå for his statistical advice and expertise. We also thank research nurse Margareta Backmän (Umeå) and project assistant Nighat Farooqi (Umeå) for collecting data and establishing invaluable contacts with the families.

REFERENCES
20. Farooqi A, Hägglöf B, Sedin G, Gothefors L, Serenius F. Chronic conditions, functional limitations and special health...
Children Plus All Nonautomobile Motorized Vehicles (Not Just All-Terrain Vehicles) Equals Injuries

Christy L. Collins, MA, Gary A. Smith, MD, DrPH, R. Dawn Comstock, PhD

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

ABSTRACT

OBJECTIVES. The goals were to describe the epidemiological features of pediatric nonautomobile motorized vehicle–related injuries sustained between 1990 and 2003 and to compare all-terrain vehicle–related injuries with other types of nonautomobile motorized vehicle–related injuries.

METHODS. An analysis of nationally representative pediatric nonautomobile motorized vehicle–related injury data from the US Consumer Product Safety Commission National Electronic Injury Surveillance System was performed.

RESULTS. Nationally, an estimated 1 203 800 children were treated in hospital emergency departments for nonautomobile motorized vehicle–related injuries between 1990 and 2003. These children had a mean age of 12.7 years (range: 1 month to 19 years), and 77.0% were male. The majority of injuries were associated with all-terrain vehicles (44.8%), 2-wheeled off-road vehicles (21.1%), and go-carts/buggies (13.7%). The most common diagnoses were contusions/abrasions (28.3%), fractures (24.2%), and lacerations (20.0%). Overall, the number of injuries increased 86% from 70 500 injuries in 1990 to 130 900 injuries in 2003. The numbers of all-terrain vehicle–related, 2-wheeled off-road vehicle–related, 2-wheeled on-road vehicle–related, and go-cart/buggy-related injuries all increased significantly from 1990 to 2003. There were greater proportions of all-terrain vehicle–associated injuries among children ≥16 years of age (48.0%) and children 12 to 15 years of age (46.6%) than among children <12 years of age (40.3%). Conversely, the proportion of other nonautomobile motorized vehicle–related injuries among children <12 years of age (47.2%) was greater than that among children 12 to 15 years of age (30.3%) and children ≥16 years of age (23.0%).

CONCLUSIONS. Although most public health and legislative attention to date has been focused on all-terrain vehicles, parents, children, and public officials should be educated about the injury risk that all types of nonautomobile motorized vehicles pose to children.
In 2005 alone, there were an estimated 40,400 all-terrain vehicle (ATV)-related injuries sustained by children <16 years of age who were treated in US emergency departments (EDs). In addition, between January 1982 and December 2005, there were ~7188 ATV-related deaths, with 30% of those fatalities occurring among children <16 years of age. More than 6.5 billion dollars are spent each year for the treatment of ATV-related injuries. Despite the startling number of ATV-related injuries and deaths and the high health care costs, there are an estimated 7 million ATVs in use in the United States. This dichotomy has led some advocates and lawmakers to call for ATV legislation, including stricter ATV use guidelines and bans on the use of ATVs by adolescents.

Multiple studies have examined pediatric ATV-related injuries. Additional studies have examined other types of nonautomobile motorized vehicle–related injuries, such as those associated with motorbikes, motorcycles, and snowmobiles. A few previous studies compared ATV-related injuries with other nonautomobile motorized vehicle–related injuries; however, those studies were limited to the patient population of one trauma center. The focus of most legislative efforts has been on ATVs, despite the fact that other nonautomobile motorized vehicles also present the risk of injury or death to children. For example, between 2001 and 2004, 23,800 pediatric off-road motorcycle-related injuries were treated in US EDs. A greater understanding of the pediatric injury risk associated with different types of nonautomobile motorized vehicles is needed to drive effective recommendations, guidelines, and legislation.

This study is the first to use a nationally representative sample to describe the epidemiological features of all pediatric nonautomobile motorized vehicle–related injuries and to compare ATV-related injuries with injuries associated with other types of nonautomobile motorized vehicles.

METHODS

Data were obtained from the US Consumer Product Safety Commission (CPSC) National Electronic Injury Surveillance System (NEISS), which collects information on individuals treated for injuries in a nationally representative stratified probability sample of 98 US hospital EDs, including 8 children’s hospitals. The NEISS data set, which is updated daily, provides patient demographic information and specific information about the injury and injury event for each patient presenting for treatment. Statistical weights provided by the CPSC are applied to the NEISS sample data for calculation of national estimates of the number of injuries.

All injuries among children ≤19 years of age that were presented to NEISS EDs between 1990 and 2003 and were identified as nonautomobile motorized vehicle–related injuries on the basis of NEISS consumer product codes were included in the analysis. The types of nonautomobile motorized vehicles included ATVs (product codes 3285, 3286, 3287, and 3296), 2-wheeled off-road vehicles (including trail bikes and dirt bikes; product codes 3258 and 5036), 2-wheeled on-road vehicles (including licensed 2-wheeled vehicles, mopeds, minibikes, and scooters/skateboards; product codes 1910, 3215, 5035, and 5042), go-carts/buggies (including go-carts and beach and dune buggies; product codes 3259 and 3288), grass/farm-related vehicles (including tractors, golf carts, and powered riding lawnmowers; product codes 1062, 1213, 1405, and 1422), water-related vehicles (including powered personal watercraft and boats; product codes 3292 and 3298), snow-related vehicles (including snowmobiles and ice/snow boating vehicles; product codes 1290 and 3247), and unspecified vehicles (including powered riding toys and motorized vehicles not otherwise classified; product codes 1330 and 1744) (NEISS does not collect data on automobile-related injuries). Injuries not associated directly with the operation of a nonautomobile motorized vehicle (such as a child injured when tripping in the garage and falling against a parked ATV or a child injured while loading a snowmobile onto a trailer) were excluded from the analyses.

Other NEISS variables of interest were the child’s age and gender, injury diagnosis, body part injured, and injury disposition. Age was divided into 3 groups on the basis of child development milestones and current ATV regulations, that is, <12 years of age, 12 to 15 years of age, and ≥16 years of age. The 26 sites of injury (ie, body part injured) were categorized into 6 body regions, including head, face, upper extremity (including shoulder, upper arm, elbow, lower arm, wrist, hand, and finger), lower extremity (including upper leg, knee, lower leg, ankle, foot, and toe), and other (including neck, trunk, and pubic region, 25%–50% of the body, as coded in the NEISS data, and >50% of the body, also as coded in the NEISS data). Injury narratives for all cases were read to categorize 2 additional variables, namely, involvement and protective equipment. Involvement was categorized as driver, passenger, on vehicle but unspecified status, or bystander. Use of protective equipment was categorized as worn, not worn, or unspecified. This study was approved by the institutional review board of the Columbus Children’s Research Institute.

Data were analyzed by using SPSS 14.0 (SPSS, Chicago, IL) with the complex samples module, with adjustment for sample weights and the stratified survey design, as recommended by the CPSC, to produce national injury estimates. Injury estimates are rounded to the nearest 100 in the text of this article. Injury rates were calculated by using annual population estimates from the US Census Bureau. Statistical analyses included the $\chi^2$ test with Yates’ correction and linear regression. Injury proportion ratios (IPRs) were calculated by using
95% confidence intervals (CIs) and P values to assess statistical significance (P values of < .05 were considered significant). For example, the calculation comparing the proportions of head injuries among ATVs and 2-wheeled off-road vehicles is as follows: IPR = (national estimated no. of ATV-related head injuries/national estimated total no. of ATV-related injuries)/(national estimated no. of 2-wheeled off-road vehicle–related head injuries/national estimated total no. of 2-wheeled off-road vehicle–related injuries).

RESULTS

Overall Injury Epidemiological Features
Nationally, an estimated 1 203 800 children ≤19 years of age were treated in US EDs for nonautomobile motorized vehicle–related injuries between 1990 and 2003. These children had a mean age of 12.7 years (SD: 4.3 years; range: 1 month to 19 years), and 77.0% were male (Table 1). Of all nonautomobile motorized vehicle–related injuries, the majority were associated with ATVs (44.8%), 2-wheeled off-road vehicles (21.1%), and go-carts/buggies (13.7%) (Table 1). Other types of nonautomobile motorized vehicles associated with injury were 2-wheeled on-road vehicles (8.6%), grass/farm-related vehicles (7.6%), snow-related vehicles (3.2%), water-related vehicles (0.2%), and unspecified vehicles (0.8%). Regarding involvement status, the child’s involvement as a driver, passenger, or bystander was unspecified for 81.8% of injuries. However, 3.4% of injured children were bystanders who were injured when they were struck by, run over by, or burned by a nonautomobile motorized vehicle. Only 7.9% of the injury narratives in the NEISS data set mentioned the absence or presence of protective gear.

The most common diagnoses for nonautomobile motorized vehicle–related injuries were contusions/abrasions (28.3%), fractures (24.2%), lacerations (20.0%), and sprains/strains (12.3%) (Table 1). The overall pediatric nonautomobile motorized vehicle–related injuries were contusions/abrasions (28.6%), lacerations (22.3%), and fractures (20.9%). Among upper extremity injuries, the most common sites were the shoulder (23.0%), wrist (20.9%), and lower arm (19.0%), and the most common diagnoses were fractures (46.7%) and contusions/abrasions (22.8%). Of the 10.4% of injuries that involved the head, the most common diagnoses were internal injuries (31.7%), lacerations (22.8%), and concussions (22.4%). The majority of facial injuries (8.3% of all injuries) were lacerations (58.2%) and contusions/abrasions (30.1%).

Although the majority (90.8%) of children were treated and released from the ED (Table 1), an estimated 69 900 children were admitted to the hospital for treatment of their injuries, and an estimated 1900 children died. The most common body regions injured among children who were admitted were the lower extremities (31.8%), head (24.1%), and upper extremities (16.9%), and the most common diagnoses were fractures (51.3%) and internal injuries (14.5%). The most common body regions injured among children who died were the head (55.1%) and trunk (13.5%). One fifth (20.5%) of injuries that resulted in death were to >50% of the body, as coded in the NEISS data for body part injured. The most common diagnoses among children who died were internal injuries (46.8%) and not stated (29.3%). Other diagnoses among these children were submersion (10.1%), fractures (9.1%), and crushing (4.7%), also as coded in the NEISS data.

Trends Over Time According to Type of Nonautomobile Motorized Vehicle
Overall, the number of nonautomobile motorized vehicle–related injuries increased significantly from 70 500 injuries in 1990 to 130 900 injuries in 2003 (P < .01). The overall pediatric nonautomobile motorized vehicle–related injuries

### Table 1

<table>
<thead>
<tr>
<th>Characteristics of Nonautomobile Motorized Vehicle–Related Injuries Treated in EDs in the United States From 1990 to 2003 (n = 1 203 846)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age, mean, y</strong></td>
</tr>
<tr>
<td><strong>Male,</strong> %</td>
</tr>
<tr>
<td><strong>Type of motorized vehicle,</strong> %</td>
</tr>
<tr>
<td>ATV</td>
</tr>
<tr>
<td>2-wheeled off-road vehicle</td>
</tr>
<tr>
<td>Go-cart/buggy</td>
</tr>
<tr>
<td>2-wheeled on-road vehicle</td>
</tr>
<tr>
<td>Grass/farm-related vehicle</td>
</tr>
<tr>
<td>Snow-related vehicle</td>
</tr>
<tr>
<td>Water-related vehicle</td>
</tr>
<tr>
<td><strong>Unspecified</strong></td>
</tr>
<tr>
<td><strong>Diagnosis,</strong> %</td>
</tr>
<tr>
<td>Contusion/abrasion</td>
</tr>
<tr>
<td>Fracture</td>
</tr>
<tr>
<td>Laceration</td>
</tr>
<tr>
<td>Sprain/strain</td>
</tr>
<tr>
<td>Concussion</td>
</tr>
<tr>
<td><strong>Body region,</strong> %</td>
</tr>
<tr>
<td>Head</td>
</tr>
<tr>
<td>Face</td>
</tr>
<tr>
<td>Upper extremity</td>
</tr>
<tr>
<td>Lower extremity</td>
</tr>
<tr>
<td>Trunk</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td><strong>Disposition,</strong> %</td>
</tr>
<tr>
<td>Released</td>
</tr>
<tr>
<td>Admitted</td>
</tr>
<tr>
<td>Transferred</td>
</tr>
<tr>
<td>Held for observation</td>
</tr>
<tr>
<td>Left against medical advice</td>
</tr>
<tr>
<td>Fatality/dead on arrival</td>
</tr>
<tr>
<td>Not recorded</td>
</tr>
</tbody>
</table>

*a Only the most common diagnoses and concussions are included.*

---

**DISPOSITION, %**

**BODY REGION, %**

**COLLINS et al**
The trend over time in the numbers of injuries according to type of nonautomobile motorized vehicle. In each year from 1990 through 2003, ATVs accounted for the greatest proportion of all motor vehicle–related injuries (40.0%–49.0% of total injuries in each year). The number of ATV-related injuries increased significantly from 1990 to 2003 (P < .01). Although fewer injuries were related to 2-wheeled off-road vehicles, the trends over time for these nonautomobile motorized vehicles were very similar to the trends for ATVs. The number of 2-wheeled off-road vehicle–related injuries increased significantly from 14,000 injuries in 1990 to 31,700 injuries in 2003 (P < .01). In addition, from 1990 through 2003, the numbers of 2-wheeled on-road vehicle–related injuries (P = .05) and go-cart/buggy-related injuries (P < .01) increased significantly.

### ATVs, Compared With Other Types of Nonautomobile Motorized Vehicles

Among ATV-related, 2-wheeled off-road vehicle–related, and other nonautomobile motorized vehicle–related injuries, the 4 most common diagnoses were contusion/abrasions, fractures, lacerations, and sprain/strains (Table 2). A greater proportion of 2-wheeled off-road vehicle–related injuries were fractures (30.9%), compared with the proportion of fractures among ATV cases (25.3%; IPR: 1.22; 95% CI: 1.15–1.30; P < .01) and other nonautomobile motorized vehicle–related cases (18.4%; IPR: 1.69; 95% CI: 1.57–1.81; P < .01). In all 3 groups of nonautomobile motorized vehicles, upper and lower extremities represented the greatest proportions of injuries (Table 2). However, the proportion of upper extremity injuries among 2-wheeled off-road vehicle–related injuries (36.9%) was greater than that among ATV cases (29.8%; IPR: 1.24; 95% CI: 1.18–1.31; P < .01) and other nonautomobile motorized vehicle cases (31.4%; IPR: 1.18; 95% CI: 1.11–1.26; P < .01). The proportion of lower extremity injuries among 2-wheeled off-road vehicle cases (37.1%) was also greater than that among ATV cases (30.3%; IPR: 1.22; 95% CI: 1.14–1.30; P < .01) and other nonautomobile motorized vehicle cases (31.2%; IPR: 1.18; 95% CI: 1.10–1.27; P < .01). Conversely, there were greater proportions of head injuries from ATVs (11.6%; IPR: 1.73; 95% CI: 1.43–2.03; P < .01) and other nonautomobile motorized vehicles (11.0%; IPR: 1.62; 95% CI: 1.40–1.88; P < .01) than from 2-wheeled off-road vehicles (6.7%). There were also greater proportions of facial injuries from ATVs (9.2%; IPR: 1.89; 95% CI: 1.58–2.25; P < .01) and other nonautomobile motorized vehicles (9.0%; IPR: 1.85; 95% CI: 1.54–2.23; P < .01) than from 2-wheeled off-road vehicles (4.9%).

The majority of patients with ATV-related injuries (88.9%), 2-wheeled off-road vehicle–related injuries (92.4%), and other nonautomobile motorized vehicle–related injuries (92.1%) were treated and released (Table 2). However, greater proportions of patients with ATV-related injuries were admitted (6.8%) or transferred (3.2%), compared with those with 2-wheeled off-road vehicle–related injuries (5.1% and 1.6%, respectively; admitted: IPR: 1.32; 95% CI: 1.07–1.65; P < .01; transferred: IPR: 1.96; 95% CI: 1.35–2.85; P < .01) or other nonautomobile motorized vehicle–related injuries (5.0% and 1.8%, respectively; admitted: IPR: 1.36; 95% CI: 1.17–1.58; P < .01; transferred: IPR: 1.71; 95% CI: 1.33–2.20; P < .01). Of the 69,900 children who were admitted, most injuries were related to ATVs (52.4%) or 2-wheeled off-road vehicles (18.6%). Among the 1900 children who died, the majority of injuries were related to ATVs (58.7%) or grass/farm-related vehicles (18.1%), which are included in the other nonautomobile motorized vehicle category in Table 2.

### Comparison of Age Groups

In general, between 1990 and 2003, the number of nonautomobile motorized vehicle–related injuries treated in US EDs increased in all 3 age categories (<12 years, 12–15 years, and ≥16 years) (Fig 2). In each year except 1997, children 12 to 15 years of age accounted for the greatest proportion of all motor vehicle–related injuries (35.0%–41.7% of injuries in each year). The number of nonautomobile motorized vehicle–related injuries sustained by children <12 years of age increased significantly from 23,000 injuries in 1990 to 44,400 injuries in 2003 (P < .01). The numbers of nonautomobile motorized vehicle–related injuries sustained by children 12 to 15 years of age and children ≥16 years of age also increased significantly between 1990 and 2003 (P < .01 for both age groups).
In all age groups, the majority of nonautomobile motorized vehicle–related injuries were associated with ATVs (Table 3). However, there were greater proportions of ATV-associated injuries among children 16 years of age (48.0%; IPR: 1.19; 95% CI: 1.11–1.28; \( P < .01 \)) and children 12 to 15 years of age (46.6%; IPR: 1.15; 95% CI: 1.10–1.21; \( P < .01 \)) than among children 12 years of age (40.3%). There were also greater proportions of injuries associated with 2-wheeled off-road vehicles sustained by children 16 years of age (29.0%; IPR: 2.33; 95% CI: 2.01–2.70; \( P < .01 \)) and children 12 to 15 years of age (23.2%; IPR: 1.85; 95% CI: 1.68–2.05; \( P < .01 \)) than by children <12 years of age (12.5%). Conversely, the proportion of other nonautomobile motorized vehicle–related injuries among children <12 years of age (47.2%) was greater than those among children 12 to 15 years of age (30.3%; IPR: 1.53; 95% CI: 1.43–1.64; \( P < .01 \)) and children ≥16 years of age (23.0%; IPR: 2.00; 95% CI: 1.80–2.21; \( P < .01 \)).

In each age category, the majority of injured children were on the vehicle with unspecified driver or passenger status (Table 3). However, children 12 years of age sustained a significantly greater proportion of nonautomobile motorized vehicle–related injuries as bystanders (5.9%) than did children 12 to 15 years of age (2.3%; IPR: 2.61; 95% CI: 2.12–3.20; \( P < .01 \)) and children ≥16 years of age (1.7%; IPR: 3.45; 95% CI: 2.66–4.47; \( P < .01 \)). Children <12 years of age also sustained significantly greater proportions of head and facial injuries (12.6% and 12.5%, respectively) than did children 12 to 15 years of age (9.9% and 5.7%, respectively; head injuries: IPR: 1.28; 95% CI: 1.16–1.42; \( P < .01 \); facial injuries: IPR: 2.20; 95% CI: 1.94–2.49; \( P < .01 \)) and children ≥16 years of age (8.2% and 6.5%, respectively; head injuries: IPR: 1.55; 95% CI: 1.35–1.77; \( P < .01 \); facial injuries: IPR: 1.92; 95% CI: 1.67–2.22; \( P < .01 \)). However, children 12 to 15 years of age sustained a significantly greater proportion of concussions (2.8%) than did children <12 years of age (1.8%; IPR: 1.31; 95% CI: 1.00–1.72; \( P = .06 \)).

### TABLE 2
Characteristics of ATV-Related, 2-Wheeled Off-Road Vehicle–Related, and Other Nonautomobile Motorized Vehicle–Related Injuries Treated in EDs in the United States From 1990 to 2003, According to Body Region, Diagnosis, and Disposition

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>ATVs (n = 539,486)</th>
<th>2-Wheeled Off-Road Vehicles (n = 253,927)</th>
<th>Other Motorized Vehicles (n = 410,433)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contusion/abrasion</td>
<td>27.5</td>
<td>24.5</td>
<td>31.9</td>
</tr>
<tr>
<td>Fracture</td>
<td>25.3</td>
<td>30.9</td>
<td>18.4</td>
</tr>
<tr>
<td>Laceration</td>
<td>19.4</td>
<td>16.9</td>
<td>22.8</td>
</tr>
<tr>
<td>Sprain/strain</td>
<td>12.6</td>
<td>13.5</td>
<td>11.3</td>
</tr>
<tr>
<td>Concussion</td>
<td>2.7</td>
<td>2.2</td>
<td>1.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Body region</th>
<th>ATVs</th>
<th>2-Wheeled Off-Road Vehicles</th>
<th>Other Motorized Vehicles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>11.6</td>
<td>6.7</td>
<td>11.0</td>
</tr>
<tr>
<td>Face</td>
<td>9.2</td>
<td>4.9</td>
<td>9.0</td>
</tr>
<tr>
<td>Upper extremity</td>
<td>29.8</td>
<td>36.9</td>
<td>31.4</td>
</tr>
<tr>
<td>Lower extremity</td>
<td>30.3</td>
<td>37.1</td>
<td>31.2</td>
</tr>
<tr>
<td>Trunk</td>
<td>10.0</td>
<td>9.2</td>
<td>8.4</td>
</tr>
<tr>
<td>Other</td>
<td>9.1</td>
<td>5.2</td>
<td>9.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Disposition</th>
<th>ATVs</th>
<th>2-Wheeled Off-Road Vehicles</th>
<th>Other Motorized Vehicles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Released</td>
<td>88.9</td>
<td>92.4</td>
<td>92.1</td>
</tr>
<tr>
<td>Admitted</td>
<td>6.8</td>
<td>5.1</td>
<td>5.0</td>
</tr>
<tr>
<td>Transferred</td>
<td>3.2</td>
<td>1.6</td>
<td>1.8</td>
</tr>
<tr>
<td>Fatality/dead on arrival</td>
<td>0.2</td>
<td>0.0</td>
<td>0.2</td>
</tr>
</tbody>
</table>

\( ^a \) Only the most common diagnoses and concussions are included.  
\( ^b \) Only the most common dispositions are included.
DISCUSSION

This study is the first to use a nationally representative sample to describe the epidemiological features of all pediatric nonautomobile motorized vehicle–related injuries and to compare ATV-related injuries with all other types of nonautomobile motorized vehicle–related injuries sustained by children. A great deal of attention has been given to the prevention of pediatric ATV-related injuries, but relatively little attention has been given toward the prevention of pediatric injuries related to other types of nonautomobile motorized vehicles. There has been a relative lack of attention toward these other nonautomobile motorized vehicle–related injuries, which collectively made up more than one half of the 1.2 million pediatric nonautomobile motorized vehicle–related injuries sustained nationally from 1990 through 2003.

Consistent with previous studies, we found that the number of pediatric ATV-related injuries treated in US EDs increased significantly from 1990 through 2003. During the same time period, the number of injuries related to other types of nonautomobile motorized vehicles also increased. In particular, although substantially fewer injuries were associated with 2-wheeled off-road vehicles, compared with ATVs, the trends over time for 2-wheeled off-road vehicle–related injuries and ATV-related injuries were very similar. In 2000, the American Academy of Pediatrics recommended that children <16 years of age should be prohibited from using 2-wheeled off-road vehicles and ATVs. With the popularity of 2-wheeled off-road vehicles continuing to increase, additional focus should be placed on the prevention of 2-wheeled off-road vehicle–related injuries as well as ATV-related injuries.

Pediatric injuries associated with other nonautomobile motorized vehicles, including 2-wheeled on-road vehicles, go-carts, buggies, grass/farm-related vehicles, snow-related vehicles, and water-related vehicles, are also of concern. Because the proportions of head and facial injuries associated with ATVs and other nonautomobile motorized vehicles were very similar, children operating or riding on any type of nonautomobile motorized vehicle should wear a helmet with facial protection. Other nonautomobile motorized vehicle–related injuries are of particular concern among children <12 years of age, because nearly one half (47.2%) of all injuries sustained by this age group were associated with nonautomobile motorized vehicles other than ATVs and 2-wheeled off-road vehicles. More specifically, 1 of 5 injuries sustained by children <12 years of age were go-cart/buggy-related and another 12.5% of injuries were grass/farm vehicle–related. Formal guidelines and recommendations should be developed for the use of all types of nonautomobile motorized vehicles, especially with respect to younger children.

Although many states have some age regulations regarding the use of ATVs, requirements vary by state, with minimal operating ages ranging from 10 years to 18 years. In this study, we found that, in almost every year from 1990 to 2003, children 12 to 15 years of age sustained the greatest proportion of nonautomobile motorized vehicle–related injuries. More than two thirds of injuries sustained by these children were ATV or 2-wheeled off-road vehicle related. Children 12 to 15 years of age also sustained a significantly greater proportion of concussions than did children <12 years and ≥16 years of age. On the basis of these findings, we recommend that specific nonautomobile motorized vehicle–

---

**TABLE 3** Characteristics of Nonautomobile Motorized Vehicle–Related Injuries Treated in EDs in the United States From 1990 to 2003, According to Age

<table>
<thead>
<tr>
<th>Proportion, %</th>
<th>&lt;12 y (n = 412,494)</th>
<th>12–15 y (n = 463,145)</th>
<th>≥16 y (n = 328,207)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>71.8</td>
<td>77.8</td>
<td>82.2</td>
</tr>
<tr>
<td>Type of motorized vehicle</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATV</td>
<td>40.3</td>
<td>46.6</td>
<td>48.0</td>
</tr>
<tr>
<td>2-wheeled off-road vehicle</td>
<td>12.5</td>
<td>23.2</td>
<td>29.0</td>
</tr>
<tr>
<td>Go-cart/buggy</td>
<td>22.0</td>
<td>11.4</td>
<td>6.7</td>
</tr>
<tr>
<td>2-wheeled on-road vehicle</td>
<td>8.0</td>
<td>10.3</td>
<td>6.8</td>
</tr>
<tr>
<td>Grass/farm-related vehicle</td>
<td>12.5</td>
<td>5.4</td>
<td>4.6</td>
</tr>
<tr>
<td>Snow-related vehicle</td>
<td>2.2</td>
<td>2.8</td>
<td>4.8</td>
</tr>
<tr>
<td>Water-related vehicle</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>Unspecified</td>
<td>2.3</td>
<td>0.1</td>
<td>0.0</td>
</tr>
<tr>
<td>Involvement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Driver</td>
<td>4.3</td>
<td>6.0</td>
<td>6.2</td>
</tr>
<tr>
<td>Passenger</td>
<td>4.2</td>
<td>2.8</td>
<td>2.1</td>
</tr>
<tr>
<td>On vehicle but unspecified</td>
<td>77.2</td>
<td>83.5</td>
<td>85.2</td>
</tr>
<tr>
<td>Bystander</td>
<td>5.9</td>
<td>2.3</td>
<td>1.7</td>
</tr>
<tr>
<td>Unspecified</td>
<td>8.4</td>
<td>5.4</td>
<td>4.8</td>
</tr>
<tr>
<td>Diagnosisa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contusion/abrasion</td>
<td>29.3</td>
<td>27.0</td>
<td>29.0</td>
</tr>
<tr>
<td>Fracture</td>
<td>21.6</td>
<td>26.9</td>
<td>23.4</td>
</tr>
<tr>
<td>Laceration</td>
<td>25.0</td>
<td>18.5</td>
<td>15.8</td>
</tr>
<tr>
<td>Sprain/strain</td>
<td>8.4</td>
<td>12.7</td>
<td>16.9</td>
</tr>
<tr>
<td>Concussion</td>
<td>1.8</td>
<td>2.8</td>
<td>2.4</td>
</tr>
<tr>
<td>Body region</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Head</td>
<td>12.6</td>
<td>9.9</td>
<td>8.2</td>
</tr>
<tr>
<td>Face</td>
<td>12.5</td>
<td>5.7</td>
<td>6.5</td>
</tr>
<tr>
<td>Upper extremity</td>
<td>28.0</td>
<td>34.0</td>
<td>33.7</td>
</tr>
<tr>
<td>Lower extremity</td>
<td>29.1</td>
<td>35.0</td>
<td>31.5</td>
</tr>
<tr>
<td>Trunk</td>
<td>8.6</td>
<td>8.1</td>
<td>11.8</td>
</tr>
<tr>
<td>Other</td>
<td>9.2</td>
<td>7.3</td>
<td>8.3</td>
</tr>
<tr>
<td>Disposition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Released</td>
<td>90.7</td>
<td>90.4</td>
<td>91.2</td>
</tr>
<tr>
<td>Admitted</td>
<td>5.2</td>
<td>6.4</td>
<td>5.7</td>
</tr>
<tr>
<td>Transferred</td>
<td>2.7</td>
<td>2.3</td>
<td>2.1</td>
</tr>
<tr>
<td>Held for observation</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>Left against medical advice</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Fatality/dead on arrival</td>
<td>0.3</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Not recorded</td>
<td>0.8</td>
<td>0.4</td>
<td>0.5</td>
</tr>
</tbody>
</table>

*Only the most common diagnoses and concussions are included.*
related injury prevention efforts should be targeted at this age group and that regulations to limit the use of nonautomobile motorized vehicles should include all children <16 years of age.

The main limitations of this study were associated with the data set. The NEISS data provide only limited additional information, of inconsistent quality and breadth, about injury events in the narratives. For example, for 81.8% of nonautomobile motorized vehicle–related injuries, the child was reported in the narrative to have been on the vehicle; however, it was not specified whether the child was a driver or a passenger, which is important for the development of targeted injury prevention programs and policies. Furthermore, only 7.9% of the injury narratives in the NEISS data set mentioned the absence or presence of protective equipment. Adding a protective equipment variable to NEISS, even just a yes/no/not applicable categorization of protective equipment use, would provide researchers with valuable data for the development of injury prevention programs and policies. Information about other factors that might influence the risk of injury, such as parental supervision or nonautomobile motorized vehicle size, was also unavailable. Because such data were not available in the NEISS data set, future studies are needed to examine the impact of these factors on all types of pediatric nonautomobile motorized vehicle–related injuries. Another limitation is that children treated in EDs may not be representative of all children who are injured during nonautomobile motorized vehicle–related activities, because less severely injured children may seek other or no medical attention. Despite these limitations, the NEISS data set provides the only nationally representative sample of nonautomobile motorized vehicle–related injuries in the United States. Although exposure-based injury risk rates could not be calculated because of a lack of denominator data, such as the number of children who actually ride nonautomobile motorized vehicles and the frequency of rides or the amount of time spent on the vehicle, comparisons of estimated numbers of injuries, analyses of trends over time, and descriptions of patterns of injuries based on this stable long-term database yielded important information.

All types of nonautomobile motorized vehicles pose a risk of injury to children. As shown by the number of injuries found in this study and in similar studies, children, especially those <16 years of age, do not have the judgment and motor skills needed to operate any type of nonautomobile motorized vehicle safely.24 On the basis of our findings, we support several safety recommendations. As recommended by the American Academy of Pediatrics, children <16 years of age should be restricted from riding 2-wheeled, 3-wheeled, and 4-wheeled off-road vehicles.24 Although this study was limited by the amount of available data on protective gear, our findings regarding injuries to the head and face support the recommendation that appropriate protective equipment, including helmets and eye protection, should always be worn when nonautomobile motorized vehicles are ridden. Because one fifth of injuries that resulted in death were to >50% of the body, future research is needed to examine the risk of fatal injury resulting from the heavy weight of the vehicle and potential for high impact.

Parents, children, and public officials should be educated about the injury risk all types of nonautomobile motorized vehicles pose to children. In addition, parents should be aware of the risk of injury that nonautomobile motorized vehicles pose to all children, including passengers and bystanders as well as drivers. Future research is needed to evaluate existing legislation for ATV and 2-wheeled vehicle use (ie, minimal requirements for operations and helmet requirements), to determine effectiveness. All states in the United States should be encouraged to adopt the most effective policies, and these policies should be expanded to include all types of nonautomobile motorized vehicles.

ACKNOWLEDGMENTS

This study was not supported by any funding agency. We thank the US CPSC and the NEISS for providing the data.

REFERENCES


IN PRAISE OF LESS PRAISE

“Readers wrote about soccer leagues that don’t keep score to avoid hurt feelings: so the kids keep score in their heads. And parents have to pay ‘trophy fees’ before sports seasons even start. Kids know these trophies are bought and not earned. Several readers sent me dialogue from the 2004 animated film The Incredibles. There’s a scene in which the superhero mom tells her son, ‘Everyone’s special!’ The boy mutters: ‘Which is another way of saying no one is.’”

Noted by JFL, MD
Injuries in Canadian Youth Ice Hockey: The Influence of Relative Age

Nick Wattie, BPHE, BSc, Stephen Cobley, BSc, MA, Alison Macpherson, PhD, Andrew Howard, MD, MSc, FRCCS, William J. Montelpare, PhD, Joseph Baker, PhD

School of Kinesiology and Health Science, York University, Toronto, Ontario, Canada; Carnegie Research Institute, Leeds Metropolitan University, West Yorkshire, United Kingdom; Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada; Division of Orthopaedic Surgery, Hospital for Sick Children, Toronto, Ontario, Canada; Population Health Sciences, Hospital for Sick Children Research Institute, Toronto, Ontario, Canada; School of Kinesiology, Lakehead University, Thunder Bay, Ontario, Canada

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

ABSTRACT

OBJECTIVE. The purpose of this study was to investigate the relationship between relative age and injury prevalence in Canadian youth ice hockey.

METHODS. In study 1, youth ice hockey–related injuries (among children 10–15 years of age) collected by the Canadian Hospitals Injury Reporting and Prevention Program between 1995 and 2002 were analyzed. The relative ages of injured children were compared across different age groups and injury characteristics (mechanism of injury and severity of injury). In study 2, injuries reported in the Hockey Canada Insurance Database were analyzed. The relative ages of injured children at different levels of play (ie, representative versus house league teams) were compared.

RESULTS. In study 1, the majority of injured players were of older relative age. However, relative age was not related to mechanism of injury or severity of injury. In study 2, ~40% of injured players at the highest level of play were relatively older, whereas only 20% to 25% of house league injured players were relatively older.

CONCLUSION. Relatively older children within ice hockey age groups are at increased risk of injury compared with their younger peers. Furthermore, the risk of injury for relatively older players is greater at more competitive levels of play. This study proposes that the relative age advantage associated with selection to Canadian youth ice hockey teams is accompanied by an increased risk of injury.
At first glance, annual age-grouping strategies in both sports and education seem to be a developmentally appropriate way to distribute children. This has been achieved by establishing “cutoff” date criteria, which ensure that children have turned a specific age before being placed in a certain grade for education or being eligible to participate on certain sports teams. For example, with a cutoff date of September 1, only children who had turned 6 years of age before September 1 of the current academic year would be allowed to enter first grade. However, there are problems with such an age-grouping strategy, as exemplified by relative age effects (RAEs).1,2 RAEs describe the potential advantages (or disadvantages) that result from age differences between peers within an annually age-grouped cohort. From the previous example, someone born shortly after September 1 would be almost 1 year older than someone born on August 30, but both would be in the same grade.

In sports, the consequences of such relative age differences suggest that relatively older members of a cohort are more likely to be selected for local junior teams3,4 and representative teams5,6 and to attain elite professional status2,7,8 which highlights participation and attainment inequality. RAEs have been identified in many sports (eg, baseball,8 ice hockey,9 swimming,10 and basketball11) and across cultural contexts (eg, in Belgium, Holland, and France,12 England,13 and Australia, Brazil, Germany, and Japan14). Regardless of alterations in cutoff dates or climatic seasonal variations, age-grouping RAEs persist.15,16 To date, physical maturational,6,17 experiential,18 and psychological16 hypotheses have been presented to account for RAE prevalence. Specifically, relative age differences, especially in prematurity stages, may create significant height and weight differences among youths participating in sports. This may confer specific physical performance advantages for relatively older players,19,20 leading to identification and selection for higher levels of competition. With the amount of time spent in sport-specific practice being closely related to attainment21 and there being variability of lived experience between relatively older and younger players, the amount of potential practice and game experience may result in a skill acquisition disadvantage for relatively younger players. Relatively younger players may not display the equivalent skills and competencies, compared with their relatively older peers. The psychological hypothesis suggests that personal perceptions of ability and competence directly influence behavior and participation.22 Positive competence is thought to reinforce and to maintain motivation for a given behavior (eg, sports participation), whereas negative perceptions, including affect, emotion, and stress, may lead to cessation. Therefore, relatively younger players may develop low perceived competence when skills or performance are not equivalent to those of their relatively older counterparts.16 Such perceptions may lead to reductions in effort, interest, and investment and ultimately to termination of sports participation.23

Separately or possibly in combination, these hypotheses suggest how age-grouping systems can lead to selection and attainment discrepancies in sports. To date, most relative age research has focused on simply identifying selection and attainment advantages associated with relative age differences. However, these may not be the only unintentional outcomes of age grouping; other implications may be evident but as yet not identified. For instance, in sports (eg, ice hockey) in which physical factors such as height, weight, speed, power, and coordination underlie performance, it is plausible that such differences may lead to discrepancies in injury incidence according to relative age. Currently, research has not examined intra-age group (ie, relative age) patterns of injury.

The purpose of the present study was to examine whether relative age was related to physical injury as a result of participation in youth ice hockey (a physical-contact, high-strategy, team sport) and whether the proportions of relative age injuries changed across youth development phases (ie, inter-age group). Furthermore, given the physical contact inherent to ice hockey and the hypothesized influence of physical size and strength on creating RAEs, a secondary purpose of the current study was to examine whether players of different relative ages were more or less likely to be injured as the result of body contact or body checking and more or less likely to incur severe injuries. This secondary purpose was of interest because previous research identified body checking as being associated with injury severity.24

Previous research on the influence of relative age found advantages exclusively for relatively older children, compared with their younger peers.16 It has been hypothesized that relatively younger players are at a physical disadvantage with respect to size and strength and that this disadvantage, compared with their older peers, results in less selection to youth ice hockey teams. We hypothesized that these maturational physical discrepancies between players of different relative ages would result in a greater proportion of injured players being relatively younger. Similarly, because relatively older players are thought to be larger and stronger than their younger peers, we hypothesized that relatively younger injured players would have incurred a greater proportion of their injuries through physically aggressive contact (ie, body checks). Because body checks have been related to more severe injuries, we also hypothesized that relatively younger players would incur a greater proportion of severe injuries.

Level of play (or competition level) was identified previously as an important covariate in relative age research, and this is particularly true of relative age research on youth ice hockey. Barnsley and Thompson17 observed that the RAE in youth hockey was more pro-
nounced at higher tiers in all age groupings. It was even observed that the lowest tiers of competition demonstrate a significant excess of relatively younger children. If level of play can influence the magnitude and existence of RAEs in a sport, then due consideration of this covariate is necessary. The purpose of the second study was to examine a secondary hockey-injury data source that contained information on level of play, in addition to information on age group and relative age.

**METHODS**

**Study 1**

**Sample**

The study sample consisted of boys 10 to 15 years of age who visited pediatric emergency departments participating in the Canadian Hospitals Injury Reporting and Prevention Program (CHIRPP). The CHIRPP is an injury surveillance program conducted in certain pediatric emergency departments. All participating hospitals in Ontario and Quebec (excluding the Children’s Hospital of Eastern Ontario) were included. All ice hockey–related injuries reported by the CHIRPP between September 1995 and August 2002 were included. CHIRPP data include demographic variables and numerous variables pertaining to injuries. Specific injury information, such as injury type and severity, were reported by the treating physician, whereas information on how the injury was sustained was provided by the patient and parents.

**Measures**

The primary independent measure was the relative age of the injured ice hockey players, as reported by the CHIRPP. Consistent with previous relative age research, players’ birth dates were grouped into monthly quartiles starting from the selection date criteria (in this case, starting on January 1 and ending on December 31 of the same calendar year). For example, quartile 1 included January, February, and March birth dates and quartile 4 included October, November, and December birth dates. Subjects born in quartile 1 were therefore relatively older than their peers born in quartile 4.

Several moderating variables were examined, to ascertain their interactions with relative age and subsequent RAEs. From 1995 to 2002 (7 ice hockey seasons), the youth ice hockey age groups dictated by Hockey Canada were Atom (10–11 years), Peewee (12–13 years), and Bantam (14–15 years). To determine whether relative age injury proportions were moderated by age groups, the proportions of relative age injuries were compared at different age groups.

Another moderating variable of interest was the mechanism factor, reported as “how the injury occurred” on the CHIRPP form. Injury was dichotomized as either resulting from a body check or not resulting from a body check. Hockey Canada defines body check- as the “physical extension of the body toward the puck carrier moving in an opposite or parallel direction.” Methods for classifying injuries as checking related or not checking related were consistent with previous use of this CHIRPP data set for examination of youth ice hockey–related injuries.

In light of the previously identified relationship between injuries resulting from body checks and the severity of injuries, a measure of injury severity was included in the analyses. Injuries were categorized as severe if the athlete was admitted to the hospital and nonsevere if the athlete was not admitted.

**Analyses**

Statistical analyses were conducted by using SPSS 15.0 for Windows (SPSS, Chicago, IL). Frequencies of relative age (ie, quartile of birth) and age group (Atom, Peewee, or Bantam) were used to describe the study sample; χ² statistics (significance of P < .05, assuming equal distribution) were used to compare the observed versus expected proportions of relative age injuries in each of the age groups (Atom, Peewee, and Bantam). Logistic regression analyses were used to examine whether relative age was associated with (1) the risk of being injured in different age groups, (2) the risk of being injured by a body check, and (3) the risk of incurring a severe injury. The aforementioned risks (odds ratios [ORs]) were calculated by comparing the proportion of quartile 1 injured players with that of quartile 4 injured players (using 95% confidence intervals [CIs]).

**Study 2**

**Sample**

The Hockey Canada National Insurance Program provides financial resources to help compensate players and families that have experienced financial loss as the result of hockey participation in Canada. The study sample consisted of hockey injuries reported to the Hockey Canada National Insurance Program between September 1998 and August 2003. Each claim made to the Hockey Canada Insurance Database (HCID) contained information on the date of injury and how the injury occurred, as well as each player’s date of birth, age group, and level of play.

**Measures**

The primary independent measure of this study was the relative age of the injured athletes. Players’ birth dates were grouped into monthly quartiles (quartile 1: January to March; quartile 2: April to June; quartile 3: July to September; quartile 4: October to December).

A moderating variable of interest included in the analyses was the age group of the injured players. The age groups included in the HCID were Peewee (12–13 years) and Bantam (14–15 years). Within each of the
age groups, there were 10 level-of-play tiers, including AAA (the highest, competitive-level tier), AA, A, B, C, CC, D, DD, E, and house league (the lowest tier). For the purpose of analysis, the “intermediate” B, C, CC, D, DD, and E tiers were combined into 1 group, to maintain disclosure confidentiality of players. Therefore, in the analyses, 5 levels of play were examined as potential moderating variables (AAA, AA, A, intermediate tiers, and house league). Because the purpose of this study was to examine how the level of play influences the proportion of relative age injuries, all cases that contained no information on the level of play were excluded from analyses.

Information on how the injury occurred was used to create a “mechanism of injury” variable. Twelve descriptors were used throughout the HCID to express how injuries occurred, of which 10 were used to create the mechanism of injury variable. Injuries that had no information on how the injured occurred, as well as injuries that were the result of a fight, were not included in the analyses. The descriptors of “blindsiding,” “checked from behind,” “collision with opponent,” and “collision on open ice” were categorized as injury resulting from a body check. Conversely, the descriptors of “collision with boards/net,” “fell on ice,” “hit by puck/stick,” and “noncontact injury” were categorized as injury that did not result from a body check.

Analyses
Statistical analyses were conducted using SPSS 15.0 for Windows. Frequencies and percentages of injuries according to relative age category (quartile of birth), age group (Peewee or Bantam), and level of play were used to describe the study sample. To test for injury asymmetries according to quartile of birth, χ² analyses were conducted; χ² tests (significance of P < .05, assuming equal distribution) were conducted for all levels of play and both age groups. Logistic regression analyses were used to examine (1) the risk of injury for those of different relative ages at different levels of play and (2) the risk associated with mechanism of injury (body checked versus not body checked) for those of different relative ages. For all logistic regression analyses, risks (ORs) were calculated by comparing the proportion of quartile 1 injured players with that of quartile 4 injured players (using 95% CIs).

RESULTS
Study 1
There were 4736 ice hockey injuries reported by the CHIRPP between 1995 and 2002. Overall, the proportion of injuries varied significantly according to age group. Of the overall sample, 18% were children in the Atom (10–11 years) age group, 37% in the Peewee (12–13 years) age group, and 46% in the Bantam (14–15 years) age group (χ² = 591.13; P < .001). The proportions of injuries according to relative age category are presented in Table 1. A RAE was found within each of the 3 age groups. Birth dates (by quartile) were distributed unevenly, with a greater-than-expected number of injuries sustained by quartile 1 players and a less-than-expected number of injuries sustained by quartile 4 players (Atom: χ² = 11.20; P = .001; Peewee: χ² = 32.57; P < .001; Bantam: χ² = 42.04; P < .001). Quartiles 2 and 3 did not differ significantly from each other or from expected values and were excluded from analyses; however, parallel analyses were conducted with those quartiles included, and the direction and significance of the observed RAEs were not affected.

Although a RAE was observed within each of the 3 age groups, regression analyses suggested that the risk of injuries according to relative age did not change from one age group to another. More specifically, on the basis of a comparison of the risk of injury for relatively older players versus relatively younger players in the Atom age group, the risk of injury for relatively older players (quartile 1) did not increase or decrease in the Peewee (OR: 1.05; 95% CI: 0.82–1.34) and Bantam (OR: 1.05; 95% CI: 0.83–1.33) age groups.

Neither older (quartile 1) nor younger (quartile 4) relative age was associated with a disproportionate risk of being injured as the result of a body check in any of the age groups (Atom: OR: 0.83; 95% CI: 0.53–1.31; Peewee: OR: 0.95; 95% CI: 0.71–1.25; Bantam: OR: 1.05; 95% CI: 0.82–1.34). Furthermore, neither older nor younger relative age was found to be a significant risk factor for incurring severe injuries in any age group (Atom: OR: 1.79; 95% CI: 0.47–6.80; Peewee: OR: 1.08; 95% CI: 0.43–2.72; Bantam: OR: 0.73; 95% CI: 0.37–1.45).

Study 2
Of the 6864 (Peewee and Bantam) injuries reported by the HCID, 5681 had information on the variables of interest in this study and were retained for additional analyses. Of the final study sample, a greater proportion of injured players were in the Bantam age group (58%) than in the Peewee age group (χ² = 144.16; P < .001).

The proportions of relative age injuries (quartile 1 versus quartile 4) for the Peewee and Bantam age groups are presented in Figs 1 and 2. The χ² analyses revealed

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Frequency and Proportion of Injuries According to Relative Age Category</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age Group</strong></td>
<td><strong>Quartile 1</strong></td>
</tr>
<tr>
<td>Atom (10–11 y)</td>
<td>221 (26.6)</td>
</tr>
<tr>
<td>Peewee (12–13 y)</td>
<td>506 (29.1)</td>
</tr>
<tr>
<td>Bantam (14–15 y)</td>
<td>645 (29.8)</td>
</tr>
</tbody>
</table>

*Birth quartile 1 was January to March; quartile 4 was October to December.*
that, within both the Peewee and Bantam age groups, RAEs existed at various levels of competition. Within the Peewee age group (Fig 1), greater proportions of injuries were incurred by players born in quartile 1 than in quartile 4 for the AAA, AA, and A levels of play ($\chi^2 = 21.05–36.16; P < .001$) but not for the intermediate level of play ($\chi^2 = 1.01; P = .31$) or the house league level of play ($\chi^2 = 2.05; P = .15$). The same was true within the Bantam age group (Fig 2); greater proportions of injured players were born in quartile 1 than in quartile 4 for the AAA, AA, and A levels of play ($\chi^2 = 10.65–73.59; P < .01$) but not for the intermediate level of play ($\chi^2 = 1.45; P = .70$) or the house league level of play ($\chi^2 = 0.01; P = .89$). As in study 1, quartile 2 and quartile 3 were excluded from analyses because they did not differ significantly from each other or from expected values.

Results from the logistic regression analysis showed that there were intra-level-of-play variations in the proportions of relative age injuries in both the Peewee and Bantam age groups (Table 2). More specifically, the magnitude of the injury RAE varied according to the level of play. For example, Bantam AAA injured players were 3.21 times more likely to be relatively older (ie, born in quartile 1, rather than in quartile 4) than were Bantam house league injured players. In both age groups, the risk of injury for relatively older players increased as the level of play increased from least competitive to most competitive. Logistic regression analyses revealed that differences in relative age (in both age groups and all levels of play) did not result in disproportionate risk of being injured as the result of a body check.

DISCUSSION

Study 1
Results of this study suggest that there are both inter-age group patterns of injury, with overall injury prevalence increasing progressively from Atom to Bantam age groups, and intra-age group patterns of injury. Within the Atom, Peewee, and Bantam age groups, greater proportions of children (26%–29%) were born in quartile 1, compared with quartile 4 (18%–20%), and thus were of older relative age. Therefore, whereas relatively older children are more likely to be selected for youth hockey teams, they are also more likely to present to hospital emergency departments with hockey-related injuries. Within the scope of relative age research, these findings are decidedly atypical; previous research in both education and sports found consistent advantages in attainment for relatively older children, compared with their relatively younger peers. Interestingly, these results suggest that the advantage in selection and attainment for relatively older players is accompanied by the disadvantage of increased proportion of injuries. The overall prevalence of injury increased progressively from Atom to Bantam age groups but the risk of injury for relatively older players did not change, which suggests that the injury disadvantage for relatively older players is stable.

Although it showed surprising results, this study did have limitations. First, the study was not population based. Although the CHIRPP has been found to be a valid indicator of overall youth injury patterns in Canada, ideally the relative age of injured children would have been compared with that of children who were not injured. Second, the nature of the data collected by the CHIRPP did not provide information on the children’s level of play (ie, competitive versus recreational/house leagues) within each age grouping. Previous research found the RAE in youth hockey to be more pronounced at higher levels of play (ie, more-competitive hockey) and nonexistent, or even reversed, at lower levels of play. Because level of play has been identified as a major covariate of the RAE, failure to control for level of play represents a major limitation of this study and may
even account for the atypical direction of the RAE observed for youth ice hockey injuries. We focused on this limitation in study 2.

Study 2

The results of study 2 suggest that the prevalence of injuries increases from younger age groups to older age groups (ie, from Peewee to Bantam) and that, within each age group, relatively older players experience a greater proportion of injuries than do their relatively younger peers. Like study 1, study 2 presented both inter-age group patterns of injury and intra-age group patterns of injury. Furthermore, as in study 1, relative age was not related to mechanism of injury (ie, being injured by a body check or not). However, the addition of level of play as a covariate in study 2 demonstrated that the relationship between relative age and youth hockey injury was more dynamic than revealed previously. The addition of level of play as a covariate suggested that the proportion and risk of injury increased for relatively older players as the level of play became more competitive (Table 2).

Like study 1, this study was limited by the fact that it was not population based. Future research would benefit from examination of both injured and noninjured players when injury RAES are considered. Although this study did control for level of play, the HCID did not provide a measure of injury severity. Although there was no relationship between relative age and mechanism of injury, as in study 1, it cannot be assumed that level of play does not interact with relative age and severity of injury.

General Observations

Both study 1 and study 2 support the notion that relatively older players at the highest levels of competition are at increased risk of injury. Although at first it might seem counterintuitive that these children, with more experience and skill, would be at greater injury risk, it is possible that they are actually playing more, playing at higher competitive levels, or both and thus are more likely to be injured because of greater exposure. Research investigating the RAE in soccer not only found a relative age advantage for selection to soccer teams but also found that relatively older players had more playing time in games, compared with their younger peers.16 Because relatively older participants are selected because of their greater physical maturity, skill, and experience, it seems reasonable that these individuals would play longer and more often. If the same disproportionate amount of playing time is present for relatively older children in Canadian youth ice hockey, then these players would have increased exposure to the inherent risks associated with playing the game. Researchers should

![FIGURE 2](https://example.com/figure2.png)

**Percentages of injuries in older (quartile 1) and younger (quartile 4) relative age categories for different levels of play in the Bantam age group. Q1 indicates quartile 1; Q4, quartile 4.**

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Level of Play</th>
<th>OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peewee (12–13 y)</td>
<td>AAA</td>
<td>3.53 (1.90–6.56)</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>1.78 (1.24–2.56)</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>1.42 (1.00–2.01)</td>
</tr>
<tr>
<td></td>
<td>B/C/D/E</td>
<td>0.93 (0.66–1.32)</td>
</tr>
<tr>
<td></td>
<td>House league</td>
<td>1.00 (referent)</td>
</tr>
<tr>
<td>Bantam (14–15 y)</td>
<td>AAA</td>
<td>3.21 (2.18–4.73)</td>
</tr>
<tr>
<td></td>
<td>AA</td>
<td>2.23 (1.58–3.14)</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>1.45 (1.03–2.06)</td>
</tr>
<tr>
<td></td>
<td>B/C/D/E</td>
<td>1.06 (0.75–1.50)</td>
</tr>
<tr>
<td></td>
<td>House league</td>
<td>1.00 (referent)</td>
</tr>
</tbody>
</table>

The risk of injury for relatively older players (quartile 1) at the different levels of play was based on comparison with the lowest level of play (house league).
consider differences in exposure time as a possible explanatory factor in this unusual disadvantage for relatively older players.

Our results suggest that the profile of RAEs is more complex than considered previously. In addition to the selection and attainment advantages noted consistently for relatively older players in a range of sports, these participants are at increased risk of injury.

**Implications for Coaches and Athletes**

The nature of the data in study 2 did not make it possible to examine the interaction between relative age, level of play, and injury severity. Such research would benefit from consideration of injury severity when relative age and level of play are examined, because this might provide a more compelling view of the implications of relative age injuries. Similarly, information regarding how relative age differences affect the length of time an injured athlete is away from play or participation may be more tangible to coaches, parents, and athletes. It is noteworthy that increased injuries have been identified as one of the reasons why, as a result of decreased fun and performance, youths drop out of sports.31

There has been a great deal of debate regarding strategies and rule changes aimed at decreasing the risk of injury in Canadian youth ice hockey.24,32 The results from this study suggest that, regardless of the strategy adopted, coaches, parents, and athletes should be cognizant of the potential factors leading to increased injury risk for relatively older ice hockey players, particularly at higher levels of competition.

**ACKNOWLEDGMENTS**

We thank the CHIRPP and its participating sites, as well as Hockey Canada, for providing the data for the current study.

**REFERENCES**

2006 Job Lewis Smith Award Acceptance Address: Is There an OWA (Other Weird Arrangement) in Your Future?

Stanley I. Fisch, MD, FAAP

Harlingen Pediatrics Associates, Harlingen, Texas

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

I am greatly honored to stand before you as the recipient of this award. To stand in the line of men and women who have been here before me and to stand before you, my colleagues, teachers and friends, is astonishing and gratifying, a truly wondrous moment. Please allow me to use this occasion to reconnect with our history and to look ahead at what lies before us.

I am particularly concerned about OWAs [other weird arrangements]. We have had many OWAs come and go. Like many attractive ideas, they have their day, then fade, with traces taken up by the next fad, the next iteration of the core idea. Here are a few:

THE QUALITY MOVEMENT IN HEALTH CARE

- Quality assurance
- Quality improvement
- Performance improvement
- Six Sigma
- Patient safety
- Reducing errors

RATIONALIZING THE SYSTEM

- Managed care
- Managed cost
- Integrated delivery systems
- Disease management
- Case management
- Pay for performance

IS COMMUNITY PEDIATRICS DESTINED TO BE AN OWA WHOSE APPEAL WILL FADE?

We have had our share of OWAs in community pediatrics:

- Community medicine
Community pediatrics
- Social medicine
- Family-centered care

In addition, several themes have entered the lexicon of community pediatrics:
- Social determinants of health
- Rights and equity
- Access to care
- Overcoming disparities
- New morbidity
- Millennial morbidity

But, I believe community pediatrics as a concept and as a discipline has a lot more durability than most OWAs because of its rich lineage, substantial content, and profound relevance to today’s environment of pediatric practice.

LINEAGE
We owe much to the work and legacies of many who have labored in the field of community pediatrics. To name a few who have influenced me, and by no means constituting an exhaustive list:

Martin Cherkasky, Victor Sidel, and Harold Wise made Montefiore Medical Center and the Albert Einstein College of Medicine [both in Bronx, NY] a fount of innovations in what was then called social medicine; from prepaid group practice to neighborhood health centers to community outreach, Montefiore was and remains an extraordinary place for practice, teaching, and research in social medicine.

George Silver, an internist like Cherkasky, Sidel, and Wise, was chief of social medicine at Montefiore from 1951 to 1965. He was very interested in the interface between public health and clinical practice. From Montefiore he moved on to the Department of Health, Education, and Welfare and had a profound influence on federal health programs for many years beyond his tenure as deputy assistant secretary. Our colleagues Fernando Guerra, Jeff Goldhagen, and Peter Simon expand the work of Dr Silver in their inspiring roles as public health officers, clinicians, and community advocates.

During 50 productive years, Robert Haggerty has done so much to set the research agenda for community pediatrics and model ways in which pediatricians can engage in their communities to bring about improvements in health for all children.1

One figure of special importance to me was Sidney Kark [Fig 1], a family physician who, with his wife, Emily, developed models of community practice that still influence us today.

Sidney and Emily Kark were South Africans. After completing their medical studies, they were sent to Durban, and in the early 1940s as part of an extraordinary social experiment established a network of primary-level clinics in the surrounding areas of the city. This work produced a stream of path-breaking innovations in the practice of social medicine. His notion of social medicine was expansive:

Social medicine may be regarded as a practice of medicine concerned with health and disease as a function of group living. It is interested in the health of people in relation to their behavior in social groups and as such is concerned with care of the individual patient as a member of a family and of other significant groups in his daily life. It is also concerned with the health of these groups as such and with that of the whole community as a community.2

Their practice was the Pholela Health Unit, a “neighborhood health center,” which in the 1960s George Silver and others used as their model for the community health center initiative of the Office of Economic Opportunity. Sidney Kark understood the health problems of his patients as community health syndromes, which embodied the complex interactions of disease states, health conditions, family, and community life. Here is Kark’s construct of a community syndrome of malnutrition, infections, and mental illness [Fig. 2].3

He created teams of health care workers (a doctor, a nurse, and a health educator), and assigned each team to a particular neighborhood and list of families. The team was made responsible for the health of the families in that neighborhood and, indeed, for the health of the entire neighborhood. Teams were asked to do systematic assessments of health care needs in their assigned catchments and to use epidemiologic methods to catalog these needs and gauge the team’s effectiveness.

Family care constituted a central point of interest in the clinical and public health practice of the [team], . . . Study of the patient’s family life situation is very often essential in understanding . . . his condition and in the consideration of the prognosis and the care program.
This objective is one which is common to all doctors in clinical practice, whether in the preventive or curative field, general practitioners or specialists.  

In 1952, as chair of Social, Preventive, and Community Medicine of Natal University Medical School in Durban, Kark initiated a clerkship in social medicine for medical students. The clerkship required several unusual commitments from his students. First, students should have a longitudinal experience, “a continuing relationship with individual patients and . . . their families.” Further, students were expected to make “a diagnosis of the state of health of a group, with particular stress on family diagnosis and on appraisal of a community’s health.” Finally, students were asked to give “consideration of the resources available for promotion of health and medical care within family and community as well as through various agencies.” By the way, this course was spread over 3 years of the medical school’s curriculum.

Over 200 community health centers, modeled after the Pholela unit, were planned for South Africa. However, with the imposition of apartheid in 1948, these plans were suspended, and eventually the 40 health centers already built were closed. The Karks left South Africa in 1958 and went briefly to the University of North Carolina. Sidney became founding chairman of the Department of Epidemiology in the School of Public Health at the University of North Carolina in Chapel Hill. After a year, the Karks and colleagues from South Africa went to Jerusalem, established the Department of Social Medicine of the Hebrew University, and resumed their work in developing and promoting social medicine.

In 1981 he published a foundational work on what he called community-oriented primary care (COPC). COPC was the conceptual link between primary care and public health. Community-oriented Karkian health care teams would utilize tools and concepts from epidemiology and behavioral and social sciences. Health promotion through education and special projects was an essential component of COPC. Appreciation for the families’ cultural identification and beliefs was emphasized. And, collaboration and coordination with other providers and agencies were essential for fulfilling these charges.

Already in the late 1940s Kark was developing methods for outcomes measurement. “Measurement of the progress of families,” he wrote, “in a particular family practice and of the community’s health as a whole was a feature of the service. . . . A special group of health recorders was trained for the purpose and in this way it became possible to measure the changing state of health of the various communities and to evaluate the [practice’s] programs.”

Pay for performance, anyone?

Kark left us many important insights. These include:

- Medicine is collaborative. The integration of curative, preventive, and promotive care requires many skills and many hands.
- The “social” is inherent in all medicine, not just primary care.
- “Cultural competence” is crucial.
- Epidemiology is an important tool in community practice: “What pathology and physiology have meant in the development of scientific diagnosis of the individual patient so epidemiology is coming to mean in the study of those processes determining the health of a group. As such, it is the foundation science of social medicine.”
- There is a vital link between community practice and public health, a theme addressed more recently by the current generation of COPC advocates.
- Teaching and research are essential components of the practice of social medicine.
• Medicine must hold firmly to an orientation toward equity and social justice.

CONTEMPORARY OWAs: THE MEDICAL HOME

The medical home represents a strong partnership between physicians and families to address the needs of children, especially children with special needs. It is interesting that our colleagues in internal medicine have become interested in the medical home. The American College of Physicians (ACP) acknowledges the work of Cal Sia and the American Academy of Pediatrics going back to the 1960s, laying the foundation for the medical-home concept. For the ACP, the advanced medical home may be a matter of survival of the discipline of general internal medicine. In a paper with the astonishing title of “The Impending Collapse of Primary Care Medicine [and Its Implications for the State of the Nation’s Health Care],” the ACP sets out the stark terrain they see lying ahead for adults with chronic illness. There will not be anywhere near enough general internists to care for the boomers now approaching retirement. And, no one has given serious, systematic thought to how and by whom the needs of our obese teenagers will be met as they roll into their 20s and 30s with chronic metabolic disorders and their complications.

An important organizing principle, the architectural plan for the medical home, and yet another OWA is the chronic care model (CCM). This has been developed by another internist, Dr Ed Wagner of Seattle. Dr Wagner has proposed a model that is universal in application: pediatricians, as well as internists and family docs, would find this a useful way to configure their practices as medical homes to meet the new morbidity (the millennial morbidity) heading our way.

The CCM has these components [Fig 3]:

● a health care team which creates and manages a registry of patients and uses community resources and partners to serve the needs identified;

● an electronic health record to ensure complete data capture and access for all who need to know;

● office workflow which permits preparation for patients’ visits, making them purposeful events meeting needs of providers and patients alike, rather than random encounters; and

● informed and engaged patients and families, engaged with informed and prepared caregivers.

The CCM is distinctly not “disease management.” Rather, it is an organization of practice (either primary or specialty care) with a much broader focus than disease management programs operated to reduce payors’ costs. The CCM may have benefits for all patients in such a practice, not only those with chronic illness. The CCM is a blueprint for the design of the medical home, an information-age neighborhood health center.

“High-quality chronic illness care is characterized by productive interactions between practice team and patients that consistently provide the assessments, support for self-management, optimization of therapy, and follow-up associated with good outcomes. These interventions do not necessarily require face-to-face visits. Ample evidence documents the effectiveness of using the computer or telephone for this purpose.”

FIGURE 3

The CCM.

AN AGENDA FOR REACHING THE NEXT STAGE: FROM [CURRENT PROCEDURAL TERMINOLOGY] 99214 TO SOMETHING BETTER . . . BEGINNING WITH PRACTICAL MATTERS

The current third-party reimbursement system is built on patient encounters that are brief, with one clinician, and for an acute illness with little long-term impact. To deal effectively with the millennial morbidities, we must design new CPT codes and descriptors that match what we aspire to do. Codes 99214 and 99215 do not represent the work of a health care team performing medical home services:

● Clinical evaluation and management

● Education

● Care coordination (not disease management)

● Advocacy

Haggerty pointed out that without financial support for the practice of community pediatrics, many pediatricians would not be able to do it: no margin, no mission. In advocating new arrangements in practice and reimbursement, the ACP wants to press the issue so that general internal medicine, which it sees evolving as the advanced medical home for adults, can distinguish itself and be accountable for its services. We must collaborate with our colleagues in internal medicine and family medicine to effect essential changes in the environment of community-based, community-oriented practice. Richard Pan makes the case in a wonderful commentary in the September Pediatrics. We have a common interest with the internists and family doctors in promoting these new OWAs.

We must make certain that both generalists and subspecialists are engaged in these collaborative efforts. And, we must connect with advanced nurse practitioners, physician assistants, social workers, health educators, and outreach workers to develop mutual understanding of roles of team members.

We must press for changes in residency training.

The Task Force on the Future of Pediatric Education II in its 2000 report placed at the top of its list of recommendations the following: “Pediatric medical education at all levels must be based on the health needs of children in the context of the family and community.”
So how do we connect with context?
I believe Sidney Kark stated it correctly more than 40 years ago:

Medical and nursing education must include a basic understanding of sociology and psychology of the standard demanded in the biological sciences, such as physiology, if clinical experience in family practice is to be fully appreciated. Understanding and knowledge of the family-life situation in its relationship to family health are integral elements of social medicine and vital to the men and women who may become family physicians and family nurses. Not only is it of significance to those who become general practitioners, but also to many who specialize in other fields of medical practice.

For residents, continuity experience must be placed at the center of their training, not left to the periphery. Well-organized and fully supported continuity experiences are essential for learning about context.

Program directors must have flexibility to be creative and distinctive in what they offer. There must be less focus on inputs and process [and] more focus on outcomes. We must press the Residency Review Committee and program directors to enhance and expand the Accreditation Council for Graduate Medical Education’s general competencies with the Community Pediatrics Training Initiative’s [CPTI] competencies [Table 1].18 And, fortunately, through the CPTI, we have excellent models of how this can be done.

We can connect better with context by enlisting community pediatricians to partner with academic health centers and offer learning opportunities to students and residents. We will soon offer the Starter Kit for Community Preceptors to orient community pediatricians to their roles as teachers and mentors.

At the institutional level, the [AAP] and the Ambulatory Pediatric Association should expand their collaboration, particularly via the Community Pediatrics Education and Training Special Interest Group on the [AAP] side and the Community Based Teaching SIG [special interest group] on the APA [Ambulatory Pediatric Association] side. There are also unrealized opportunities for collaboration in informatics, curriculum, advocacy, and faculty development.

### TABLE 1 Competencies for Community Pediatrics

<table>
<thead>
<tr>
<th>Competency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delivery of culturally effective care</td>
</tr>
<tr>
<td>Child advocacy</td>
</tr>
<tr>
<td>Medical home</td>
</tr>
<tr>
<td>Special populations</td>
</tr>
<tr>
<td>The pediatrician as a consultant, partner, and collaborative leader</td>
</tr>
<tr>
<td>Educational and child care settings</td>
</tr>
<tr>
<td>Community and public health</td>
</tr>
<tr>
<td>Research and scholarship</td>
</tr>
</tbody>
</table>

**REFERENCES**

3. Kark JD, Abramson JH. Sidney Kark’s contributions to epide-

**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.
Determinants of Outcome After Head Cooling for Neonatal Encephalopathy

John M. Keogh, FRANZCOG

Obstetrics and Gynaecology, Hornsby Hospital, Hornsby, New South Wales, Australia

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

If the inhabitants of Olympus were to set 10 tasks for Hercules in perinatology, then the prevention and limitation of neurologic injury would be high on the list, along with the prevention of preeclampsia, prematurity, postpartum hemorrhage, and the impact of poverty on pregnancy outcome.

The Cool Cap (Olympic Medical, Seattle, WA)1 Trial is a step along this path. The role of this technology is as-yet uncertain, its long-term benefits are unclear, and its safety profile is still incompletely defined. However, given the frequency of adverse outcomes in this group (>60% in the whole cohort), the promising outcomes of the original randomized trials strongly suggest that cooling will be an important part of our armamentarium in the years ahead.

The latest article2 from the group presents an exploratory posthoc analysis of the data from the Cool-Cap Trial for the purpose of hypothesis generation to inform additional studies of factors that may influence the outcome of treatment.

The original study (and, therefore, the conclusions drawn from it) was, to some extent, held hostage by the difficulties inherent in the prospective diagnosis of hypoxic-ischemic encephalopathy (HIE). This subgroup of infants with newborn encephalopathy is of particular interest because of the association with acute intrapartum problems. A pathology that is in evolution at the time of birth may be reversible. Unfortunately, the criteria used to demonstrate a hypoxic-ischemic etiology as set out in the study’s inclusion criteria (ie, Apgar score of ≤5 at 10 minutes, need for resuscitation at 10 minutes, cord pH < 7.00, or base deficit of ≥16 mmol/L) were particularly nonspecific. The most specific characteristic, the cord gases, were available, even with this research subgroup, in only a minority of cases. Perhaps this accounts for the high prevalence figure for HIE quoted of 2 to 3 in 1000. In the population-based case-control study by Badawi et al,1,4 newborn encephalopathy (of all etiologies) occurred in 3.75 in 1000 births, only 29% of which were thought to have possible intrapartum hypoxia when using very inclusive criteria. Pure HIE was calculated to occur in 1.6 in 10 000 term births.

Of course, the difficulty is that by the time the underlying etiology is identified, the chance for intervention in the subgroup that may respond is lost. As yet, no reliable test exists to identify those who have had a severe insult in the hours before birth. Therefore, because potentially reversible pathologies are diluted by those in which the injury is distant to the birth, studies that seek to accurately identify subgroups that may benefit from or be harmed by treatment need to be large and are difficult to mount.

Even with these reservations, this analysis raises some interesting possibilities. The authors confirmed the prognostic significance of grade of encephalopathy and severity of electroencephalographic abnormality as predictors of outcome. They demonstrated that selective head cooling was effective in the management of both moderate and severe encephalopathy. They also identified an intriguing association between birth weight and out-

Abbreviation: HIE, hypoxic-ischemic encephalopathy

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.2007-0602
doi:10.1542/peds.2007-0602

Accepted for publication Mar 1, 2007

Address correspondence to John M. Keogh, FRANZCOG, Obstetrics and Gynaecology, Hornsby Hospital, Palmerton Road, Hornsby, New South Wales 2077, Australia. E-mail: keoghj1@optushome.com.au

come. Although the division of birth weight into those below and above the 25th centile seems a little arbitrary, given the sample size, it is probably not unreasonable. In the analysis presented, larger infants were both more likely to have an adverse outcome, in terms of death or severe disability at 18 months of age, and more likely to show a greater therapeutic response to treatment. The authors put forward several possible explanations for these results. However, it remains an unexpected finding. In the Badawi et al study of newborn encephalopathy, infants >4 kg had the lowest incidence of newborn encephalopathy of any group, whereas birth weight of <3rd centile was the risk factor with the greatest magnitude of effect.

So what other possible explanations may help explain this association with birth weight? It is possible that, compared with growth-restricted fat-depleted fetuses, larger infants may respond to episodes of profound hypoxia by metabolizing fat with the production of ketone bodies and free radicals. Larger infants in the face of hypoxia still have to meet the metabolic demands of the large placenta and body, thus diverting precious oxygen and blood flow from central organs. Interestingly, the only group of large infants who were at increased risk of newborn encephalopathy in the Badawi et al study were those infants >4 kg who had an acute onset of severe preeclampsia with its associated placental ischaemia but no time to decrease their metabolic demands by growth restricting. These infants were more likely to develop signs of intrapartum hypoxia and were at a 14-fold increased risk of newborn encephalopathy. Larger infants also have greater head volumes and greater volumes of cerebral blood flow. This may lead to differential cooling of superficial and deep cerebral structures, which might confer a benefit. Finally, because larger infants, in general, seem to be protected from newborn encephalopathy, it is quite possible that encephalopathy in such an infant is indicative of a more severe insult, especially when it is thought to follow an acute insult. This may help to explain the trend to worse outcomes with increasing birth weight, and it would be interesting to see the analysis for those infants >90th centile for birth weight.

The findings in relation to fever are, I believe, of great interest given the link between intrapartum fever and adverse neurologic outcome (such as cerebral palsy) demonstrated by Grether and Nelson, among others. Although it is not yet known whether control of fever changes the outcome, it is a tempting proposition. In the cooling group there were 11 infants with a temperature >38°C either before cooling or at rewarming. Nine of these infants had an adverse outcome compared with 50 of 97 in the cooling group who did not experience fever (odds ratio: 4.2 [95% confidence interval: 0.97–18]). In this analysis, the contribution of birth weight dominated that of fever; however, given the constancy of this association, it seems likely that fever is an independent risk factor for adverse outcome.

Although there was no statistically significant difference, there was a trend toward poorer outcomes in smaller infants who were cooled compared with those in the control group: 24 of 39 in the intervention group compared with 14 of 33 in the controls (odds ratio: 2.17 [95% CI: 0.85–5.54]). This may well be from chance alone; however, given the small numbers in this subgroup analysis, it is quite possible that it represents a true effect. Additional studies will need to address this question. Certainly no benefit has been shown in these smaller infants. In future studies, it may also be of value to differentiate between infants who are small for gestational age and those who are growth restricted. The pathways that lead to growth restriction may well have affected the ability of the neonate to cope with acute injury and not necessarily in a negative manner.

Cooling, either selective or whole-body, is one of the few therapies shown in randomized studies to have a definite neuroprotective effect. It is in its infancy, but it is here to stay. In time we will know more about how to optimally apply it and to whom. It is an exciting time to be alive.

REFERENCES

Pediatricians and Medical Malpractice

Steven M. Donn, MD, FAAP,a,b Gary McAbee, DO, JD, FAAP, FACOP, FCLMb,c

aDivision of Neonatal-Perinatal Medicine, Department of Pediatrics, University of Michigan Health System, Ann Arbor, Michigan; bAmerican Academy of Pediatrics Committee on Medical Liability and Risk Management; cDepartment of Pediatrics (Neurology), Robert Wood Johnson School of Medicine, Camden, New Jersey

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

THE RECENT Pediatrics articles by Carroll and Buddenbaum1 and Kain and Caldwell-Andrews2 analyzed pediatric medical malpractice claims from 2 important sources: the Physician Insurers Association of America data-sharing project and the National Practitioner Data Bank. The authors themselves note the limitations of these data, yet there are important aspects of the litigation process that can be gleaned from this work.

Pediatricians should not avoid perusing these types of articles. Because the law is based on precedent, an important risk-management technique is for the practitioner to be aware of the reasons for being sued (eg, errors in diagnosis), the medical conditions that place them at the highest risk for being sued, and the amounts of indemnity being paid so that adequate malpractice insurance is maintained. This information is important in implementing effective risk-management techniques to not only minimize the risk of being sued but also to result in improved patient safety. Anesthesiologists were one of the first groups to analyze their high rate of malpractice lawsuits by focusing on patient safety while assessing why lawsuits occurred. This resulted in the institution of strict internal standards and organized initiatives, with a dramatic reduction in number of lawsuits and indemnity paid (20th lowest of 28 specialties in median indemnity paid).3

These data also demonstrate issues that are relevant to effective tort reform. Over the 20-year period (1985–2005), 67% to 80% (mean: 72%) of lawsuits were either dropped or settled without payment. Although the defense costs were lower than in cases in which payments were made, the total costs involved in these unsuccessful lawsuits are staggering. Better efforts to monitor these cases (eg, oversight of expert witnesses) might result in significant savings. Furthermore, it takes ~5 years to settle a pediatric malpractice case, and 54% of cases settle for less than $100 000. For these reasons, alternative dispute-resolution systems should be studied. One apparently successful approach bypasses the litigation process by instituting procedures with which mistakes are acknowledged, apologies are made, and settlements are paid internally.4

Only 6% of cases involved “no medical misadventure,” which suggests that frivolous lawsuits uncommonly result in payment to the plaintiff. This confirms what others5 have found, and this information should be disseminated to the legal community.

The enlightening news from the Physician Insurers Association of America study is that only 5% of cases go to trial (and verdict), and only 20% result in a plaintiff verdict. This should assuage some of the emotional burden of pediatrician defendants, because most lawsuits are not likely to become publicized, and details can be maintained within the practice or hospital setting.

Unfortunately, medical malpractice lawsuits have become part of professional life. Lawsuits are not uncommon in business, and they are a burden that will continue to occur as part of the business end of medical practice. Our tort system is outcomes based and, as such, practitioners will continue to be sued. Tort reform is only a partial solution. Perhaps a better approach would be focusing on patient safety and quality improvement while simultaneously using sound risk-management.

References

principles; communicating effectively among providers, patients and families; and improving medical chart documentation.

REFERENCES
4. Boothman RC. Apologies and a strong defense at the University of Michigan Health System. Physician Exec. 2006;32:7–10

HOW “LEAGUE TABLES” DEVOLVED INTO WALL STREET’S BIG JOKE

“The face of Wall Street changed back on July 14, 1982, the first time this newspaper published rankings of investment banks’ merger advising and underwriting prowess. These league tables, as they later became known, treated banking as if it were a perpetual pennant race, a welcome jolt for an industry that for decades had avoided competing. Twenty-five years later, the league tables have become a beloved institution for journalists and bankers hoping to extract both meaning and bonuses from the frenzy of street life. But the real meaning might be a farce. The tables have become home to the most petty and wheedling impulses of the industry’s most-respected institutions, which are rabid about staying high in the rankings. . . . ‘Clients laugh about league tables,’ says Herald L. Ritch, formerly mergers chief at Donaldson, Lufkin & Jenrette and now president of Sagent Advisors. ‘Everyone comes in and says they’re No. 1 in the world in everything. It’s a stupid way to run a business, and because it makes a bunch of grown-ups look silly to their clients, it’s not helpful.’ . . . Still, it is about time the farce came to an end.”

Editor’s Note: Hospital and college “rankings” are just as silly!
Noted by JFL, MD
COMMENTARY

Bilirubin Toxicity to Human Erythrocytes: A More Sanguine View

Antony F. McDonagh, PhD

Division of Gastroenterology and the Liver Center, University of California, San Francisco, California

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

In 1962 Watson observed that a small fraction of the bilirubin in blood from jaundiced newborns is associated with red cells. A few years later, Cheung et al showed that red cells undergo time- and concentration-dependent changes in morphology and metabolism, with eventual hemolysis, when incubated with bilirubin in buffer, particularly at concentrations of >340 μM (20 mg/dL). (In retrospect, it seems highly likely that the bilirubin solutions to which the cells were exposed were supersaturated and may have contained aggregated forms of bilirubin.) The bilirubin-induced morphologic changes were illustrated by striking photographs similar to those published some 30 years later, without reference to the earlier work. Subsequently, Brito et al have confirmed and extended these observations and recently reviewed the toxicity of bilirubin to red cells (again without alluding to the seminal earlier work). For the neonatologist, an important question is whether bilirubin-induced hemolysis is clinically significant in neonatal jaundice or in other unconjugated hyperbilirubinemias. The review implies that it is. Yet, they do not.

1. Gunn rats and patients with severe Crigler-Najjar syndrome have lifelong unconjugated hyperbilirubinemia yet no abnormal hemolysis. In fact, the human disorder was originally described as a “non-hemolytic jaundice,” and patients have been described who showed no abnormal erythrocyte morphology, reticulocytosis, or other evidence of hemolysis despite prolonged exposure to high plasma bilirubin concentrations. Comparisons between nonjaundiced and jaundiced rats with congenital unconjugated bilirubinemia (Gunn rats) showed no difference in hematocrit values in male rats and only slightly lowered values in jaundiced female rats. No difference in the osmotic fragility of red cells from jaundiced and nonjaundiced animals was detected. If anything, bilirubin is protective.

2. Because hemolysis leads to increased bilirubin formation, bilirubin concentrations in Gunn rats and patients with Crigler-Najjar would be expected to steadily increase with time if bilirubin-induced hemolysis is significant. Yet, they do not.

3. As evidence for a toxic effect of bilirubin on red cells in jaundiced infants, Brito et al reported a greater proportion of morphologically abnormal red cells in jaundiced neonates compared with “healthy” newborns. However, in that study, the healthy blood was umbilical blood, whereas the jaundiced blood was venous blood taken 2 to 4 days after birth, which makes for a questionable comparison. Furthermore, the association between jaundice and increased numbers of abnormal cells does not necessarily imply that the former led to the latter. Brown et al observed misshapen red cells (dubbed pyknocytes) in neonatal blood but found no association between their number and bilirubin concentrations. When they exposed adult red cells to serum from a severely jaundiced
infant with pyknocytosis, changes in morphology were not seen, which suggests that the original pyknocytosis was not caused by bilirubin. Pyknocytosis has been observed mainly in jaundiced infants, and most authors have concluded that it is caused by an endogenous extracellular factor(s) in blood. However, investigations have failed to unequivocally implicate bilirubin as that factor. Whether bilirubin is the cause or end result of the pyknocytosis, or even both, remains an unanswered question.

4. Vest and Grieder found no relationship between bilirubin levels and the rate of red cell destruction in neonates or that prolonged hyperbilirubinemia correlated with short erythrocyte survival. Their observations are consistent with earlier observations that had also indicated that plasma bilirubin does not cause the mechanical fragility of infant red cells.

5. Intravenous injection of unconjugated bilirubin was used extensively in the past as a liver-function test, and bilirubin has been injected into infants and adults in many experimental studies without notable hemolysis. The hypothesis that bilirubin might induce erythrocyte hemolysis is based mainly on studies of the cytotoxicity of bilirubin in vitro. In most of these studies, the pigment was presented to cells as a solution in buffer or culture media in the presence or absence of albumin or other proteins. Because crystalline bilirubin does not dissolve in these solvents, the solutions are invariably made by adding a concentrated solution of bilirubin in strong alkali at a pH of ~13 (or sometimes in dimethyl sulfoxide or ethanolamine) to the medium, often followed by neutralization with strong acid. It is generally assumed that the structure and properties of bilirubin in the final solutions are identical to those of bilirubin as it is presented to erythrocytes or other cells in vivo. However, that assumption may not always be valid because of the ability of bilirubin to form albumin complexes of different stoichiometry and stereochemistry; its conformational lability and ability to form hydrogen-bonded complexes; and its well-known tendency to aggregate. Absent from most in vitro studies of bilirubin toxicity are data on the structure, molecularity, and stability of bilirubin in the media used, in both the presence and absence of cells. The preponderance of studies in which significant toxic effects have been observed, including those using displacing agents, were performed under nonphysiologic conditions in which aggregates of bilirubin may have been present. There is, however, no evidence for the presence of bilirubin aggregates in neonatal or Crigler-Najjar red cells or plasma, even at high bilirubin concentrations, and evidence that they are unlikely to occur. The striking bilirubin-induced erythrocyte damage graphed in Fig 1 of the recent review becomes less impressive from a clinical point of view when it is recognized that it was obtained by exposure of the cells to concentrations of bilirubin that were orders of magnitude greater than the free bilirubin concentrations to which red cells are exposed in jaundiced infants. Also, it would not have occurred if physiologic concentrations of albumin had been present in the medium. In addition, because the aqueous bilirubin solutions to which the red cells were exposed were most likely supersaturated, their admixture with cells may well have seeded nucleation, aggregation, and precipitation of pigment, resulting in nonphysiologic multiphasic mixtures. Few studies of the toxic effects of bilirubin on red cells have been performed in whole blood. Some have been performed with washed red cells reconstituted to a hematocrit level of 10%. It seems not to have been recognized in those studies that the total load of bilirubin presented to the cells for the same plasma concentration of bilirubin is much greater at a hematocrit level of 10% than at the normal ~45%. Petrich et al reported that bilirubin, added to heparinized blood from an adult at a concentration of only 70 to 80 μM (~4–5 mg/dL), induced hemolysis after 4 hours at 37°C. However, that lone observation is inconsistent with other observations on red cells in vitro, in which no hemolysis was observed at even higher bilirubin concentrations in the presence of serum albumin. Red cells from just 1 adult were used in the studies, and the possible interference of bilirubin in the spectroscopic method used to estimate hemolysis, a phenomenon noted by Watson, was not checked.

From the published clinical and in vitro evidence, with the exception of the article noted above, it is hard to escape the conclusion that bilirubin-induced hemolysis is insignificant in vivo even at plasma bilirubin concentrations associated with kernicterus or at levels greater than those at which therapeutic intervention is considered mandatory. There is even evidence that bilirubin can strengthen the red cell membrane and protect red cells from oxidative damage that could lead to hemolysis. Of course, it is hard to substantiate a negative. Documentation of mild hemolysis in newborns is difficult, and even a small degree of hemolysis, insufficient to have much effect on hematocrit measurements, could markedly augment the circulating bilirubin pool. Necheles et al and, more recently, Maisels and Kring have shown that increased heme catabolism is important in the development of hyperbilirubinemia in normal infants in the first few days after birth. The cause of this is unknown and merits additional investigation, but it would be premature to ascribe it to bilirubin-induced hemolysis.

The hypothesis that binding of bilirubin might predispose red cells of neonates to hemolysis or phagocytosis is certainly plausible and might seem to be supported by some in vitro studies. It is unquestionable that bilirubin
can partition into cell membranes, modifying their properties. However, rash extrapolations from in vitro studies to real life have previously muddied the waters of bilirubin metabolism and led, for example, to suggestions that photodegradation of bilirubin is the most important pathway in phototherapy and that photohemolysis might be a significant adverse effect of the treatment, suggestions now thought to be incorrect. It seems to have gone unnoticed that bilirubin has the characteristics of a promiscuous inhibitor. Promiscuous inhibitors are compounds that lead to false-positive results in the high-throughput screening of potential drugs in vitro. Like bilirubin, promiscuous inhibitors are often hydrophobic molecules with a tendency to aggregate, and aggregation is often the cause of their nonphysiologic promiscuous in vitro inhibition. Although the toxic effects of bilirubin on erythrocytes and other cells in vitro may be scientifically interesting with respect to the chemical properties of bilirubin and nonspecific effects of lipophiles and aggregated molecules on membranes, stronger evidence than that recently presented or currently available is needed to support the hypothesis that bilirubin-induced damage to red cells is ever sufficient to cause clinical concern.

ACKNOWLEDGMENT
This work was supported by National Institutes of Health grant DK-26307 (to Dr McDonagh).

REFERENCES
45. Maisels MJ, Kring EA. Does intensive phototherapy produce hemolysis in newborns of 35 or more weeks gestation? J Perinatol. 2006;26:498–500
Hypnosis as a Therapeutic Tool in Pediatrics

Haleh Saadat, MD, FAAP, Zeev N. Kain, MD, FAAP

Departments of Anesthesiology, Pediatrics, and Child Psychiatry, Yale University School of Medicine, New Haven, Connecticut

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

Complementary and alternative medicine (CAM) therapies are achieving considerable increase in popularity and recognition in both adult and pediatric populations. A national survey conducted in 2004 indicated that relaxation techniques, guided imagery and hypnosis, are among the most popular mind-body therapies in adults. In 1997, 43% of Americans spent $27 billion out of pocket on 1 or more alternative therapies. The frequency of using CAM for children is increasing as well, and a 2001 survey indicated that 20% to 30% of pediatric patients used 1 or more CAM therapies. These rates are much higher (30%–70%) among children with chronic and recurrent conditions such as cancer, asthma, rheumatoid arthritis, migraine headache, and cystic fibrosis. With the increased popularity of CAM in children, pediatricians are confronted with parental demands and questions related to the integration of CAM in patient management. As such, pediatricians should be aware of all available CAM modalities including hypnosis.

Hypnosis is one of the very first ancient CAM interventions and is defined as “a natural state of focused concentration coupled with a relative suspension of peripheral awareness.” This modality can be dated back several thousand years to the Greeks, Egyptians, and Persians. Milton Erickson and Ernest Hilgard were among the first investigators in the United States to undertake a modern, systematic approach to hypnosis, and the American Medical Association acknowledged hypnosis as a valuable tool in medical treatment in 1958. A National Institutes of Health Technology Assessment Panel report in 1996 judged hypnosis to be a viable and effective intervention for alleviating pain with cancer and other chronic pain.

Traditionally, a hypnotic experience consists of 3 components: (1) absorption, focused concentration; (2) dissociation, relative suspension of peripheral environment; and (3) suggestibility, a communication indicating that an individual will experience a particular response toward the goal of a therapy. Hypnosis is a particularly suitable intervention for children because, in general, children are more susceptible to hypnosis than adults. This increased susceptibility has been attributed to children’s enhanced ability and willingness to become absorbed in fantasy, play, and imagination. Indeed, clinical hypnosis has been applied as an effective adjunct in the management of a variety of pediatric disorders.

Several studies have reported the use of hypnosis in the management of painful surgical and medical procedures and postoperative pain. Hypnosis was also used both as a solo technique and an adjunct to analgesic medications for the management of acute painful conditions such as burns and fractures in pediatric emergency settings. Children who suffer from chronic pain conditions such as recurrent abdominal pain, migraine headache, and sickle cell disease have shown significant benefit from the use of hypnosis in hospital settings. Holden reviewed 31 studies of treatments for children with chronic headache and found good evidence for the efficacy of relaxation and self-hypnosis in reducing pain.

Several studies have also shown improvement with use of hypnosis in children and adolescents who suffered from behavioral conditions such as trichotillomania, thumb-sucking, chronic dyspnea, Tourette syndrome.
drome, enuresis, and dysphasia. Indeed, Anbar and Geisler found that 75% of a group of children who were taught self-hypnosis showed significant improvement in symptoms such as habitual cough, hyperventilation, shortness of breath and sighing, and vocal cord dysfunction. Several clinical reports have also consistently documented the clinical effectiveness of hypnosis in managing symptoms and improving quality of life in a population of children with asthma and cystic fibrosis. Finally, case reports have shown that hypnosis can be used for the treatment of some habitual disorders such as sleep disturbances, night terrors, swallowing problems, and nocturnal enuresis as well as some dermatologic condition such as atopic dermatitis, chronic eczema, and viral warts. Taken together, we recommend the use of hypnosis as an adjunct for the treatment of procedural pain and anxiety, phobias, sickle cell disease, and a number of chronic pain disorders such as headache and abdominal pain. This recommendation is supported by recent Cochrane reviews on this topic. The effectiveness of hypnosis for the management of other disorders such as asthma, sleep disturbances, and certain dermatologic conditions is unclear at the current time, and more data from randomized, controlled trials are needed. Pediatricians should note that introducing the concept of self-hypnosis to children early in the course of a chronic disease is advantageous, because it would give them a sense of control and mastery. Parents should also get involved early in the hypnosis process, because parental conceptions about hypnosis may either impede or assist a child’s therapy. To select a suitable hypnosis technique to a child, several factors such as developmental age, child’s condition, motivation, interest, and abilities should be considered. Indeed, more randomized, controlled trials that use validated outcome measures are still needed to address issues related to the optimal timing and length of hypnotic interventions.

ACKNOWLEDGMENT
This work was supported in part by National Institutes of Health, National Institute of Child Health and Human Development, grant R01HD37007-02.

REFERENCES

6. Integration of behavioral and relaxation approaches into the treatment of chronic pain and insomnia. NIH Technology Assessment Panel on Integration of Behavioral and Relaxation Approaches Into the Treatment of Chronic Pain and Insomnia. JAMA. 1996;276:313–318
27. Anbar RD. Self-hypnosis for anxiety associated with severe


---

**DOCTORS ASSAIL UNITEDHEALTH’S THREAT OF FINES**

“A new UnitedHealth Group Inc policy that threatens to fine doctors for referring patients to out-of-network laboratories for tests is mushrooming into a bitter dispute between the health-insurance giant and many of the 520,000 physicians in its networks nationwide. Most health plans are designed so their members pay more when they go to an out-of-network doctor or take a non-preferred medication. But the financial sanctions, which UnitedHealth has yet to impose, mark the first time a physician could be fined by a health insurer if he or she directs a patient to seek out-of-network care or testing, the American Medical Association says. The threats stem from a 10-year deal that UnitedHealth struck late last year with Laboratory Corp of America Holdings to become its national in-network laboratory. With 28.5 million health-plan members and growing, UnitedHealth has been using its heft more and more in recent years to negotiate cut-rate fees with doctors, drug makers and other suppliers.”

Fuhrmans V. *Wall Street Journal.* April 10, 2007

Noted by JFL, MD
Counting Children With Disability in Low-Income Countries: Enhancing Prevention, Promoting Child Development, and Investing in Economic Well-being

Michael E. Msall, MD, Dennis P. Hogan, PhD

Section of Developmental and Behavioral Pediatrics, Kennedy Center and Institute of Molecular Pediatric Sciences, Pritzker School of Medicine, Comer and LaRabida Children's Hospitals, University of Chicago, Chicago, Illinois; Population Studies and Training Center, Brown University, Providence, Rhode Island

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

There has been increased recognition that strategies to address women’s and children’s health require comprehensive approaches of reproductive, prenatal, neonatal, and both early- and middle-childhood public health and educational interventions.1–3 Several preventive themes have been highlighted, including prenatal nutrition, attendants at delivery, antiviral treatment for those with HIV, tetanus prophylaxis, breastfeeding during the first year, aggressive immunization campaigns, prevention of malnutrition and vitamin and trace-element deficiencies (folate, vitamin A, vitamin D, iron, zinc, iodine), oral rehydration, and community safety.4 The scope of this problem is daunting: 4 million of the 130 million infants born each year around the world die during the first 4 weeks of life.5 In addition, there are ~6 million preventable child (≤5 years old) deaths each year in developing countries.6 Major associations with these neonatal deaths include preterm birth, severe infections, and asphyxia, which in aggregate contribute to 80% of these deaths.5,7

The historic assumption has been that if one sequentially applies advances in preventive health, then the presence of child disability will be substantially reduced. In the preschool years, this has happened with reductions of motor disability from polio and cerebral palsy from iodine deficiency; deafness from rubella, measles, mumps, and bacterial meningitis; blindness from gonococcus and vitamin A deficiency; and mental retardation from vaccine-preventable encephalitis and meningitis.8,9 In the United States, medical, community, and public health advances have resulted in 98% of children entering kindergarten without neurodevelopmental disability.10 However, in developing countries in south Asia and sub-Saharan Africa, 200 million children under 5 years of age fail to reach their cognitive potential because of poverty, poor health and nutrition, and suboptimal home environments.11 These disadvantaged children are likely to do poorly in school and subsequently have low incomes, high fertility, and difficulty meeting the health and developmental needs of their own children, which results, in turn, in intergenerational transmission of poverty and compromised developmental potential.12

Because of these challenges, the article by Maulik and Darmstadt in the supplement to this month’s Pediatrics is important.13 The authors undertook a comprehensive meta-analysis to understand current knowledge on child disability in low-income (less than $875 gross national income) and middle-income ($875–$3465 gross national income) countries. What else did they discover?

First, despite the large impact on child health, family life, and economics, research in childhood disability has been woefully inadequate. Several reasons are readily apparent. They include the difficulty in counting children when there is limited public health infrastructure or community-based preventive pediatric systems. In addition, this situation is compounded if there are gaps in educational access and no formal arrangements exist...
for collaboration between health and education for evaluating children who are blind, deaf, mobility challenged, unable to follow directions, or unable to learn to read and calculate.14

Second, the Ten Questions survey was the most commonly used screening tool.15 This instrument was designed to identify children (in any culture) with sensory, motor, seizure, or severe communicative or intellectual disabilities. Although there has been success with this screening tool in several developing countries, use of the Ten Questions survey in other countries, especially with the goal of detecting mild-to-moderate intellectual disability, is far from ideal.16 A key requirement here is what might be done to promote stimulation and development in early childhood and have measures of communication and learning competencies that reflect the diverse cultural contexts of childhood. In this context, the emphasis of the Ten Questions survey on mobility, lifting, self-care, communicating, seeing, hearing, and following directions is quite good.17,18 The key need in survey research and developmental surveillance evaluation is to understand and frame indicators for social roles at key ages across elementary and secondary school experiences.19 In a developmental perspective, prevention is not a 1-age or 1-stage undertaking. In this respect, measuring functioning and activity is a promising approach and would benefit from the World Health Organization’s International Classification of Functioning (ICF) model.20

In the ICF model, a child’s health and well-being are described in terms of 4 components: (1) body structures, (2) body functions, (3) activities, and (4) participation. Body structures are anatomic parts of the body such as organs and limbs, as well as structures of the nervous, visual, auditory, and musculoskeletal systems. Body functions are the physiologic functions of body systems, including psychological functions such as being attentive, remembering, and thinking. Activities are tasks performed by children and include walking, climbing, feeding, dressing, toileting, bathing, grooming, communicating, and socially interacting with peers and adults. Participation means involvement in community life, such as playing with peers, helping with chores, and attending family activities such as visiting relatives, attending religious services, or going on errands. The ICF model also accounts for contextual factors in a child’s life, including environmental facilitators and environmental barriers as well as personal factors. Environmental facilitators include information, transportation, education accessibility, and comprehensive health services. Environmental barriers include limited health services, negative attitudes of others, lack of legal protections, and discriminatory practices. Personal factors include age, gender, interests, and sense of self-efficacy; these factors can be facilitators or barriers. Table 1 illustrates application of the ICF model for 3 children with preschool disability in developing countries. One can understand from this model community-based strategies for primary, secondary, and tertiary prevention.

Third, diagnostic advances for hearing screening using transient otoacoustic emissions and automated auditory brainstem response21 had not been systematically implemented. This could be combined with nutrition and community-wide growth, health, and developmental surveillance at key preschool and school ages.22

Fourth, in a model randomized clinical trial of service delivery for mothers of children with cerebral palsy, a comparison of health advice and play, distance training and mother-child group activities were undertaken.23 Both the distance-training package and the mother-child group activities were beneficial in improving maternal knowledge about disability-related services, reducing

### TABLE 1  ICF Model Scenarios in Preschool Children With Physical Disability in Developing Countries

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Definition</th>
<th>Girl, 3 y</th>
<th>Boy, 5 y</th>
<th>Girl, 4 y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pathophysiology</td>
<td>Molecular/cellular mechanisms</td>
<td>Burns to face, chest, and arms</td>
<td>Left cerebrovascular accident after severe dehydration</td>
<td>Hearing loss after cerebral malaria</td>
</tr>
<tr>
<td>Body structures and body functions</td>
<td>Organ structure/function</td>
<td>Unable to lift arm above head or to extend elbows</td>
<td>Hemiplegia; adaptive delays</td>
<td>Speech delays; 50-db hearing loss</td>
</tr>
<tr>
<td>Activity (functional) strengths</td>
<td>Ability to perform essential activities: feed, dress, toilet, walk, talk</td>
<td>Runs well; memorizes songs and stories</td>
<td>Walks; very strong with left hand</td>
<td>Speech understood 50% of time</td>
</tr>
<tr>
<td>Activity (functional) limitations</td>
<td>Difficulty in performing essential activities</td>
<td>Unable to carry water from well</td>
<td>Cannot run; has lost animals during chores</td>
<td>Speech not understood by peers</td>
</tr>
<tr>
<td>Participation</td>
<td>Involvement in community roles typical of peers</td>
<td>Goes to church/mosque</td>
<td>Helps in meal preparation</td>
<td>Loves to make baskets; herds animals</td>
</tr>
<tr>
<td>Participation restrictions</td>
<td>Difficulty in assuming roles typical of peers</td>
<td>Because of stigma, unable to leave hut for school</td>
<td>No soccer because he is considered “lame” by peers</td>
<td>No hearing aides available</td>
</tr>
<tr>
<td>Contextual factors: environmental facilitators or barriers</td>
<td>Attitudinal, legal, policy, and architectural facilitators</td>
<td>Has younger brother, minister at church encourages singing</td>
<td>Audiologist is available in capital city (500 miles away); home village is not on bus route</td>
<td></td>
</tr>
</tbody>
</table>
maternal stress, and improving interactions between parent and child. Thus, this pilot study demonstrated how one might link isolated populations via use of information, training, and community networks.

Fifth, protein-energy malnutrition manifested by stunting and iodine deficiency resulting in cretinism or hypothyroidism were major preventable causes of intellectual disability. In this context, it is critically important to understand rural and urban roles whereby other etiologies of intellectual disability are recognized.24

Sixth, there is increasing recognition of having cooperation and systems integration between health, social, and educational systems.25–28 It is in this context that interventions such as providing safe water, preventing burns and unintentional injuries, preventing food insecurity, and promoting parents and communities as teachers can be used as opportunities to implement systematic developmental surveillance and early childhood supports.29

Seventh, combining school vaccinations with screening of nutritional, vision, hearing, dental, and developmental status is not routine.

Eighth, the problems of stigma are real. This holds both for visible impairments (motor, blindness, sensory) and less obvious impairments (hearing, intellectual).

Ninth, gender disadvantage and discrimination against girls are also important. These issues bring us full circle to the high payoff of increasing maternal educational attainment and health literacy.

Tenth, information about severity of disability across different age groups and longitudinal studies that examine factors that promote child functioning and participation, as well as family well-being, have not taken place.

As expected, these substantial gaps in epidemiology make our prevention strategies more difficult. However, there are several points worth considering for population research:

1. What media strategies can be used to promote parents’ awareness of development? In this respect, educating parents about feeding, immunization, injury prevention, and indicators of illness can be coupled with educating them about child development.

2. How might health and education systems collaborate in preschool years and in understanding how to promote basic literacy and numeracy?

3. What survey methodologies at what ages might be useful for linking indicators of maternal health, infant health, child health, developmental status, and family well-being?

Finally, if we are to go beyond 1 disorder (eg, measles, HIV, tuberculosis) or 1 strategy (eg, oral rehydration, prenatal vitamins, vaccine preventable disorders) at a time, what key regional priorities might be implemented for enhancing prenatal, neonatal, and early childhood health and the environments of women and children? By working together to improve the status of women and children, our interventions will be an investment that enhances human capital and economic development, optimizes health and functioning, and has high payoffs in all of our futures.

ACKNOWLEDGMENT

This article was supported in part by the Brown University Office of the Vice President for Research Seed Fund Award entitled “Exceptional Children—Exceptional Challenges: Developing an Interdisciplinary Multinational Project for Studying Work-Family Dilemmas Among Parents Raising Children With Disabilities.”

REFERENCES


PHAGE-TYPING: THEN AND NOW

“A particularly virulent strain of *Staphylococcus*, named ‘80/81,’ was discovered in Australia in 1954 and very soon after in Canada, and was soon pandemic throughout hospitals and communities worldwide. This pandemic was worrisome because it occurred in the very early days of the antibiotic era. The outbreaks of nursery infections represented one of the first times that phage typers had been brought together in the battle against *Staphylococcus*. The epidemics not only brought worldwide attention to the problem of hospital cross-infection, but also brought recognition to phage typing as a useful tool for understanding the epidemiology of antibiotic-resistant staphylococcal infections. Phage typers took advantage of their time in the spotlight not only to ‘unravel the chains of infection’ but also to provide practical solutions designed to halt the spread of infection. Locally, phage typers and bacteriologists used their epidemiologic findings to assert their expertise in the hospital, to suggest administrative changes in the management of postnatal care, and to assure themselves roles on newly created infection-control committees. Incidences of infection caused by 80/81 fell drastically in the early 1960s; however, by this time hospital bacteriologists had already cemented their roles as infection-control experts, a role they continue to fulfill to this day.”


Noted by James W. Kendig, MD
Pay-for-Performance in Pediatrics: Proceed With Caution

Alyna T. Chien, MD, MS*, R. Adams Dudley, MD, MBA

*Sections of General Pediatrics and Community Health Sciences, Department of Pediatrics, University of Chicago, Chicago, Illinois; £Divisions of Pulmonary and Critical Care and Institute for Health Policy Studies, University of California, San Francisco, California

Financial Disclosure: Dr Dudley’s work on this article was supported by a Robert Wood Johnson Foundation Investigator Award in Health Policy Research.

In response to overwhelming evidence of significant quality problems within adult and pediatric health care, pay-for-performance programs have proliferated rapidly in adult care settings and are beginning to spread into pediatrics.1–4 Outpatient pediatric health care is being targeted by performance incentives in all 11 of the state Medicaid programs that currently use performance-incentive strategies and 33 of the 93 performance-incentive programs listed in the Leapfrog Compendium (the largest publicly available listing of performance-incentives programs in the country).7,8

We recognize that the current payment system contributes to our problems with quality, and we agree with the cautionary tone and measured approach suggested by Profit et al9 in the May 2007 issue of Pediatrics when considering whether performance incentives, in the form of pay-for-performance and/or public reporting, should be implemented to promote the quality of care provided by NICUs.

Because these programs require tremendous effort on the part of a wide variety of stakeholders (employers, health plans, health care organizations, and physicians), it is important to consider whether they are worth the effort. Current evidence indicates that performance-incentive strategies may only be modestly effective,10–12 are not necessarily connected to better outcomes,13,14 and can yield undesirable unintended consequences.15–20

We stress 3 general and 2 pediatric-specific issues for those considering the use of this strategy in pediatrics. The purpose of these cautionary points is to make sure that physicians, program designers, and policy makers are aware that there are risks to using performance incentives in health care and that certain performance-incentive tactics developed for adult health care will not translate well to pediatrics.

UNINTENDED CONSEQUENCES OF PERFORMANCE INCENTIVES

Recently, 3 sets of authors systematically reviewed the empirical evidence regarding performance-incentive programs in health care by using similar search and inclusion criteria.10,11,17 Although the majority of this literature assessed whether performance incentives yield their intended consequences, a small but significant handful of studies also evaluated performance incentives for their unintended effects.

1. Performance incentives can improve documentation without changing underlying quality. Within the small but growing literature on performance incentives, 2 studies indicated that performance incentives improve documentation without changing the underlying quality of care.21,22

2. Performance incentives can merely reward those already doing well. As part of the evaluation of a prototypical performance incentive program implemented in California and the Pacific Northwest, 1 study demonstrated that the vast majority of the $3.4 million in financial incentives paid to medical groups went to those who had higher baseline performance and improved the least.20

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.2007-1158
doi:10.1542/peds.2007-1158

Accepted for publication Apr 27, 2007
Address correspondence to Alyna T. Chien, MD, MS, 5841 S Maryland Avenue, MC 2007, Chicago, IL 60637. E-mail: alyna.chien@yahoo.com

3. Performance incentives can alter how willing physicians and/or health care organizations are to care for minorities and the medically complicated. An evaluation of the coronary artery bypass graft report-card effort in New York State found that black and Hispanic patients received coronary artery bypass grafts less often than their white counterparts after public reporting began.19 Additionally, a study of patients being treated for substance abuse in Maine showed that patients with the most severe substance abuse problems were less likely to be treated after Maine’s Office of Substance Abuse introduced financial incentives for improving abstinence, increasing employability, and reducing family and legal problems.18

Given the recent explosion of interest in the use of performance incentives and the substantial evidence that quality is poor under the current payment system, it is reasonable to consider whether this strategy will be helpful for improving the quality of pediatric health care. The inherent risks and challenges, which are enhanced in pediatrics, make it important to think carefully about alternatives to existing performance strategies. Researchers, physicians, and policy makers must think creatively about interventions that foster providers’ natural sense of altruism. Developing methods that support intrinsic motivation may prove to be more fruitful than performance incentives in guaranteeing long-term and sustainable improvements to our health care system.31

**ISSUES SPECIFIC TO PEDIATRICS**

The performance incentive programs that are proliferating in adult health care emphasize rewarding disease-specific processes of health care that are connected to better outcomes (eg, β blockers after acute myocardial infarction). Program designers have focused on evidence-based disease-specific processes of care for common adult conditions because providers presumably control these processes better than health care outcomes (which depend on numerous factors outside a provider’s control, such as patient preference and adherence).23–26 This general strategy faces critical challenges in pediatrics for 2 basic reasons:

1. The low prevalence of disease in pediatrics enlarges the sample-size problem in performance-incentive programs. Limited sample sizes at the provider level are already an issue for programs that target common adult conditions.15 This problem would be dramatically magnified in pediatrics. Children with condition-specific health care needs are a fraction of the total child population that is already one quarter that of the adult population.27 Those who implement pediatric-focused incentive strategies will need to pay even greater attention to methods of aggregating measures across conditions or to developing performance measures that reflect more general processes of health care (eg, measures that reflect patient-centeredness or care coordination).

2. The paucity of evidence-based quality-of-care metrics poses a greater risk of setting standards of care that are not connected to outcomes. It also increases the reliance on consensus-driven guidelines and poses a greater risk of setting standards that are not objective.26–30 Pediatric strategies will need to exert an even greater effort to make sure that goals are meaningful, realistic, and achievable by providers working in a wide range of settings. Special care should be taken to include solo and/or small group providers and those who work in less-resourced rural or urban settings.

**REFERENCES**


14. Werner RM, Bradlow ET. Relationship between Medicare’s


POLICY STATEMENT

Hepatitis A Vaccine Recommendations

Committee on Infectious Diseases

ABSTRACT
Since licensure in 1995 of a hepatitis A vaccine, the Centers for Disease Control and Prevention and the American Academy of Pediatrics have been implementing an incremental hepatitis A immunization strategy for children. In 1996, children living in populations with the highest rates of disease were targeted for immunization, and in 1999 the program was expanded to immunization of children 2 years and older living in states and counties with rates of hepatitis A that historically have been higher than the national average. The 1999 program has been successful; the current rate of hepatitis A is the lowest ever reported in the United States. Regional, ethnic, and racial differences in the incidence of hepatitis A have been eliminated. The incidence of hepatitis A in adults in immunizing states has decreased significantly, suggesting a strong herd-immunity effect associated with immunization. In 2005, the US Food and Drug Administration changed the youngest approved age of administration of hepatitis A vaccine from 24 to 12 months of age, which facilitated incorporation of the vaccine into the recommended childhood immunization schedule. As the next step in the implementation of the incremental vaccine immunization strategy, the American Academy of Pediatrics now recommends routine administration of a Food and Drug Administration–licensed hepatitis A vaccine to all children 12 to 23 months of age in all states according to a Centers for Disease Control and Prevention–approved immunization schedule. Available data suggest that hepatitis A vaccine can be coadministered with other childhood vaccines without decreasing immunogenicity. Hepatitis A vaccines have proven to be extremely safe. In prelicensure clinical trials of both Havrix (GlaxoSmithKline, Rixensart, Belgium) and Vaqta (Merck & Co Inc, Whitehouse Station, NJ), adverse events were uncommon and mild when they occurred, with resolution typically in less than 1 day. Hepatitis A vaccine is contraindicated in people with a history of severe allergic reaction to a previous dose of hepatitis A vaccine or to a vaccine component. Because the hepatitis A vaccine is an inactivated product, no special precautions are needed for administration to people who are immunocompromised. No data exist about administration of the hepatitis A vaccine to pregnant women, but because it is not a live vaccine, the risk to mother and fetus should be extremely low to nonexistent.

BACKGROUND AND RATIONALE FOR RECOMMENDATIONS
The purpose of this statement is to provide the rationale and recommendations for universal administration of hepatitis A vaccine to children living in the United States. The rationale for implementation of universal immunization is based on several considerations. For the 15 years before availability of hepatitis A vaccines
(1980–1995), approximately 30,000 cases of symptomatic hepatitis A infections (disease) were reported annually in the United States. Because of underreporting and the large number of asymptomatic infections in young children, the actual number of cases was projected to be nearly 300,000 per year in the United States. The US Food and Drug Administration (FDA) licensed the first inactivated hepatitis A vaccine (Havrix [GlaxoSmithKline, Rixensart, Belgium]) in 1995 and a second product (Vaqta [Merck & Co Inc, Whitehouse Station, NJ]) in 1996. Initial licensure limited use to people 2 years and older. With the availability of a hepatitis A vaccine, the infection became one of the most common vaccine-preventable infections in the United States.

In 1996, the Centers for Disease Control and Prevention (CDC) and American Academy of Pediatrics provided guidance for the use of hepatitis A vaccine. As part of an incremental strategy, the vaccine was recommended for use in people in specific high-risk groups, including children 2 years and older who live in defined and circumscribed communities with high endemic rates or periodic outbreaks, people with chronic liver disease, men who have sex with men, illicit drug abusers, and people with occupational hazards that put them at increased risk of acquiring hepatitis A. Despite implementation of the initial part of the strategy, hepatitis A remained one of the most frequently reported vaccine-preventable diseases, with more than 23,000 cases reported in 1996 (Fig 1). In 1999, the CDC expanded the hepatitis A immunization program to include immunization of children who live in states, counties, and communities with rates of hepatitis A consistently above the national average between 1987 and 1997. Eleven states with rates at least twice the national average (≥20 cases per 100,000) were advised to “recommend” immunization of children at 2 years of age, and an additional 6 states (with 10–20 cases per 100,000) were advised to “consider” immunization of children.

Coincident with implementation of the 1999 recommendation, hepatitis A rates dropped to the lowest rates ever recorded in the United States and in 2003 were 76% lower than the rates before immunization was initiated in 1996 (Fig 1). Before hepatitis A immunization, incidence was highest in younger children and in the American Indian/Alaska Native and Hispanic communities (Figs 2 and 3). However, since 1999, these age, ethnic, and regional differences in incidence have nearly been eliminated (Figs 2 and 3). By 2003–2004, rates of hepatitis A were highest among men 18 to 39 years of age who were residing in the states that were not advised in 1999 to immunize against hepatitis A virus (nonimmunizing states).

An additional finding associated with the implementation of immunization of children against hepatitis A was a significant decrease in the incidence of disease in adults, suggesting a strong herd-immunity effect of the immunization program. Similar findings have been reported from Israel, where within 2 years of initiation of routine immunization of children 18 to 24 months of age against hepatitis A, there was a 90% reduction in hepatitis A disease in adults throughout the country. These data suggest that focusing on routine administration of hepatitis A vaccine in young children will have a significant effect on disease incidence in the rest of the population.

The success of the interim 1999 strategy has created an opportunity to consider universal immunization of infants in the United States against hepatitis A virus. As of 2005, the FDA approved use of the hepatitis A vaccine in children as young as 12 months of age, allowing for its incorporation into the recommended early childhood immunization schedule. This approval was followed by the recent CDC recommendation for routine use of hepatitis A vaccine in all children 12 months of age and older regardless of their state of residence. Incorporation of the vaccine into the routine childhood immuni-
zation schedule also is aided by the finding that coadministra-
tion of hepatitis A vaccines with other routinely administered
immunizations has not been associated with impairment of vaccine-induced immunity. In addi-
tion, the equalization of disease rates among regions as well as among ethnic and age groups across the United States precludes sustainability of a vaccine program based on the rationale used to implement the interim strategy. Extending the program to routine immuniza-
tion of infants nationwide should result in further lowering of disease incidence in the country and possibly could lead to an environment for the eventual elimination of indigenous hepatitis A infection in the United States.

FIGURE 2

FIGURE 3

Epidemiology

Incidence and Prevalence
Hepatitis A virus has a worldwide distribution, although prevalence of infection varies considerably on the basis of hygiene and sanitation conditions. In areas with overcrowding, limited access to clean water, and inadequate sewage systems, hepatitis A infection occurs almost universally in people early in life. Because most young children who acquire hepatitis A are asymptomatic, disease rates in highly endemic areas of the world are low. Although seronegative adults in such areas of the world are at high risk of infection and disease, outbreaks are unusual because of the high prevalence of antibody to
hepatitis A virus in the population.5 High endemicity patterns also can be seen in geographic regions or ethnic groups within developed countries, including the United States before this decade.10–12

Historically, hepatitis A disease incidence has been cyclic in nature. In developed countries with temperate climates, incidence has commonly peaked every 10 to 15 years. Dramatic decreases in hepatitis A virus infection rates during the decade before and after licensure of the first hepatitis A vaccine have dampened this epidemic pattern significantly in the United States.5,13 In the pre-vaccine era, hepatitis A infection rates were highest among young children and American Indian/Alaska Native and Hispanic individuals; these differences virtually have disappeared as of 2006.5 Success of the hepatitis A immunization program has resulted in the virtual elimination of age, ethnic, racial, and regional differences in the incidence of hepatitis A infection in the past decade.5,14

Mode of Transmission

Humans, great apes, and some species of monkeys can be infected with hepatitis A virus. The primary source of hepatitis A for human transmission is person-to-person spread through the fecal-oral route. On rare occasions, hepatitis A infection has been transmitted by transfusion of blood or blood products collected from donors during the viremic phase of infection.15,16 Since 2002, nucleic acid amplification tests, such as the polymerase chain reaction (PCR) assay, have been applied to the screening of source plasma used for the manufacture of plasma-derived products.17

Transmission generally is limited to close contacts, and hepatitis A rarely is spread by casual interactions. Spread of hepatitis A within families is common, with disease occurring more commonly in older family members after being introduced into the household by an asymptptomatically infected young child.12 In child care center outbreaks, contact with feces and subsequent personal contact are important means by which transmission occurs, and cases have occurred in child care center workers and household members of children who attend the center.18,19 Foodborne hepatitis A transmission can occur from food that is contaminated during preparation by an infected food handler (foods not cooked after handling, such as salads and sandwiches) or during growing or processing (eg, produce), but this mode of transmission accounts for a relatively small proportion of reported hepatitis A cases in the United States.20 Waterborne outbreaks are rare in developed countries with adequate sanitation systems. Approximately half of the people with sporadic, community-acquired hepatitis A infection have no known source of infection.12,13,21

Stools from a hepatitis A virus–infected person are most infectious from approximately 14 to 21 days before to approximately 8 days after onset of jaundice.22 Hepatitis A RNA has been reported to be detectable in stool by polymerase chain reaction assay for up to 3 months after the acute illness;23 and children can shed hepatitis A for up to 10 weeks after onset of clinical illness.24 Although hepatitis A virus is not excreted chronically, clinical relapses may occur in 10% to 15% of patients and may be associated with recurrence of excretion of the virus in stool.25,26 Hepatitis A virus can be detected in the serum through the period of jaundice and liver enzyme elevation, which is consistent with the possible transmission of the infection by the bloodborne route.15

Hepatitis A is the most important vaccine-preventable disease for travelers. The risk of hepatitis A is 4 to 30 cases per 100 000 months of stay in an area with endemic hepatitis A for travelers who are not immunized against hepatitis A.27 In 2003, international travel was the source of hepatitis A for more than 25% of cases among children younger than 15 years. Although often not perceived as international travel by either the parents or the child’s physician, children returning from visiting family members who live in areas with endemic hepatitis A is not an uncommon source of infection among cases reported in the United States.28 Spread of hepatitis A virus in child care settings has occurred because of exposure to children who acquired hepatitis A after visiting in the countries of their parent’s birth.

CLINICAL MANIFESTATIONS OF DISEASE

Hepatitis A is an RNA virus in the Picornaviridae family. Hepatitis A virus infects the liver; the infection may be either icteric or anicteric. The likelihood of icteric (clinically apparent) disease is related inversely to the age of the person acquiring hepatitis A. In children younger than 6 years, more than 90% of hepatitis A infections are asymptomatic. In contrast, more than two thirds of older children and adults will develop jaundice after hepatitis A infection.29 These statistics explain why hepatitis A outbreaks in child care settings frequently are detected for the first time when adult contacts become jaundiced.30

Hepatitis A virus is resistant to acid, which allows for passage through the stomach to the lower intestine. After an average incubation period of 28 days (range: 15–50 days),31 infected people often experience vague and nonspecific symptoms. One of the first symptoms for which medical attention frequently is sought is dark urine, which usually is preceded by a 1- to 7-day mild prodromal illness that can include anorexia, malaise, fever, nausea, and vomiting.32,33 Within a few days of the onset of bilirubinuria, feces become clay colored, and sclera, skin, and mucous membranes become jaundiced. Hepatomegaly can be noted on physical examination. Discoloration of the stool resolves within 2 to 3 weeks, which frequently indicates resolution of disease. Pruritus occurs uncommonly. Duration of illness is variable, but most patients are significantly better within 3 to 4
weeks, including resolution of elevated hepatocellular enzyme concentrations. Among women in the United States, pregnancy is not a risk factor for more severe hepatitis A virus infections. Although transmission to the fetus is unusual, there have been 2 case reports in which mothers developed hepatitis A during the first trimester of pregnancy and their infants subsequently developed meconium peritonitis.34,35 The risk of transmission from a woman who develops hepatitis A in the third trimester of pregnancy to the infant seems to be low.36 Infants infected through this means typically are asymptomatic, but an outbreak among hospital staff related to the exposure to such an infant has been reported.37

The pathologic effects of hepatitis A are limited to the liver. As hepatitis A replicates in liver cells, virions are shed from infected hepatocytes into the hepatic sinusoids and the bile canaliculi, where they pass into the intestine and are excreted in feces. Peak infectivity occurs during the 2 weeks before onset of jaundice or serum elevation of liver enzymes.38 Viremia occurs soon after infection is acquired and persists through the period of elevated hepatocellular enzyme concentrations, but blood viral concentrations are much lower than those that occur in the stool.39

COMPLICATIONS

Approximately 10% to 15% of patients with illness attributable to hepatitis A have relapsing disease lasting up to 6 months, and approximately 20% of these people have multiple relapses.25,40 Hepatitis A virus can be detected in stool of some patients during the relapse.25,26 Even with relapsing disease, overall outcomes are very good.26,41 The clinical, laboratory, and pathologic findings in people with prolonged jaundice are associated with cholestatic hepatitis. A short course of rapidly tapered corticosteroids can reduce symptoms and hasten resolution of disease.42

Hepatitis A infection rarely results in fatalities. Before the recent success with hepatitis A immunization in the United States, there were approximately 100 deaths from hepatitis A viral infections each year in the United States. Reported case fatality rates for hepatitis A viral infections range from 0.01% to 2%. Fulminant hepatitis A viral infection is characterized by increasing severity of jaundice, deterioration in liver function, coagulation problems, and encephalopathy. Fulminant disease is more common among people older than 50 years43 and patients with chronic liver disease, including chronic hepatitis B or hepatitis C infections.44–46 Notably, serious and even fatal hepatitis A virus infection can occur in children,47,48 albeit less commonly than in people with these other risk factors. Spontaneous recovery occurs in 30% to 60% of people with fulminant hepatitis A virus disease, with survivors regaining full liver function. Prognosis is influenced by age, clotting-factor levels, stage of coma, and presence of renal disease.

VACCINE

Description

Hepatitis A vaccines licensed in the United States are inactivated, whole-cell virus vaccines that are produced from hepatitis A virus grown in human diploid fibroblast cells. There are 2 single-antigen vaccines, Vaqta49 and Havrix,50 and a combined hepatitis A/hepatitis B vaccine, Twinrix (GlaxoSmithKline). Once hepatitis A virus is adapted to growth in cell culture, it becomes attenuated. The purified virus is then formalin inactivated and adsorbed to aluminum hydroxide.51 Havrix and Twinrix have 2-phenoxyethanol added as a preservative, whereas Vaqta is preservative free. All hepatitis A vaccine preparations are administered intramuscularly. No vaccine containing hepatitis A licensed in the United States has ever contained thimerosal.

Vaccine activity in Havrix is referenced to a standard by using an enzyme-linked immunosorbent assay and is expressed, therefore, in terms of enzyme-linked immunosorbent assay units (ELU). Vaqta antigen content is expressed as units (U) of the hepatitis A antigen. The pediatric/adolescent (12 months to 18 years) dose of Havrix is 0.5 mL and contains approximately 720 ELU of hepatitis A antigen, and the 1 mL adult formulation contains approximately 1440 ELU of hepatitis A antigen.

Vaqta also is supplied in 2 formulations, one for use in children 12 months to 18 years of age and another for use in individuals 19 years of age and older. Twinrix contains a combination of hepatitis A antigen (720 ELU) and hepatitis B antigen (20 μg) and is administered as a 3-dose series on a 0-, 1-, and 6-month schedule. In the United States, Twinrix is only licensed for administration to people 18 years and older. After completion of the 3-dose Twinrix series, immunogenicity to hepatitis A and B is equivalent to immunogenicity of people who received single-antigen vaccines administered separately according to standard schedules.52

Immunogenicity

Within 1 month of receiving a first dose of hepatitis A vaccine, 97% of children and adolescents and 95% of adults developed protective concentrations of antibody, with the second dose resulting in virtually 100% of individuals being protected against the infection.53 Although data are limited, the vaccine is less immunogenic in patients with chronic liver disease, immunocompromised people, transplant recipients, and elderly individuals. Because of the high rate of seroconversion in healthy children and the insensitivity of the standardly available assays, testing for antibodies after immunization is not recommended.

Data regarding immunologic response when hepatitis A vaccines are administered concomitantly with other
routinely administered immunizations of childhood are limited. However, available data indicate that simultaneous administration of hepatitis A vaccine with diphtheria and tetanus toxoids and acellular pertussis (DTaP), poliovirus (oral and inactivated), *Haemophilus influenzae* type b (Hib), hepatitis B, or measles-mumps-rubella (MMR) vaccines did not affect immunogenicity or reactogenicity.54 No data are available regarding simultaneous administration of hepatitis A vaccine and pneumococcal conjugate vaccine (Prevnar [Wyeth Pharmaceuticals, Madison, NJ]), but there is no reason to assume that there will be an interaction between the vaccines.

Although vaccines containing hepatitis A effectively stimulate antibody production, the antibody concentrations achieved after immunization are 10 to 100 times less than concentrations that occur after natural infection.55 In addition, many commercially available tests are not sufficiently sensitive to detect the presence of antibodies against hepatitis A virus elicited by the vaccine.56 Thus, people who are immunized against hepatitis A virus may be protected against the infection but be antibody-negative according to standard assays.

**Efficacy and Effectiveness**

Two large trials have been conducted to evaluate the efficacy of hepatitis A vaccine in children. One trial, conducted in Thailand, enrolled more than 38 000 children aged 1 to 16 years who were randomly assigned to receive 2 doses separated by 1 month of either hepatitis A vaccine (Havrix, 360 ELU per dose) or hepatitis B vaccine.57 Efficacy was calculated on development of antibodies against hepatitis A more than 21 days after receipt of vaccine. Ninety-seven percent of children developed a protective titer within 1 month of immunization, and the efficacy over a 1-year period of observation after immunization was calculated at 94% (95% confidence interval: 79%–99%).57 The other trial was conducted in 1037 children aged 2 to 16 years who were living in an area of upstate New York with historically sustained high rates of transmission of hepatitis A.58 Study participants were immunized with 1 dose of Vaqta, and over the period of observation, vaccine efficacy was calculated to be 100% (the lower 95% confidence limit was 87%).58 Although long-term measurement of vaccine efficacy is needed, mathematical models predict that protective concentrations of antibody will persist more than 25 years after completion of the recommended 2-dose vaccine series.59

Postlicensure effectiveness of the hepatitis A vaccine has been shown to be similar to reported efficacy. From 1996–2000, hepatitis A vaccine was provided free of charge to children who were living in Butte County, California.60 Of the 45 000 children eligible for the program, approximately 30 000 (66%) received 1 dose and 17 600 children received 2 doses of hepatitis A vaccine. During the 5 years of surveillance, overall hepatitis rates dropped 94% in the county, and vaccine effectiveness was calculated at 98%.60 Similarly, in Catalonia, Spain, cases of hepatitis A decreased from 10.3 per 100 000 in children 11 to 14 years of age before routine hepatitis A immunization to 1.8 cases per 100 000 after implementation of a hepatitis A immunization program, resulting in an effectiveness rate of 97%.61

**Safety**

Hepatitis A vaccines have been proven to be extremely safe. In prelicensure clinical trials of both Havrix and Vaqta, adverse events were uncommon and mild when they occurred, with resolution typically occurring in less than 1 day.57 The most common adverse events, reported in 10% to 15% of subjects, were pain at the site of injection, redness, and swelling. No serious adverse events have been definitively associated with either Vaqta or Havrix. In a study of more than 38 000 Thai children who received Havrix as part of an efficacy trial, no serious adverse events were reported.57 Since licensure in 1995, millions of doses of hepatitis A vaccine have been administered, and no significant adverse events have been associated with either of the hepatitis A vaccines (Beth Bell, MD, CDC, personal communication, 2006). A postmarketing study for Vaqta was performed in a large health maintenance organization population in Northern California. During an 18-month period, patients were observed for emergency department and clinic use in the month after receipt of hepatitis A vaccine. More than 49 000 doses of hepatitis A vaccine were administered (15 000 to children younger than 18 years), and no serious adverse events were noted.62 The only vaccine-related adverse event that occurred more commonly after administration of hepatitis A vaccine was mild diarrhea in immunized adults. A summary of adverse events reported through the Vaccine Adverse Event Reporting System (VAERS) showed that 871 adverse events occurred in temporal relationship to receipt of hepatitis A vaccines. However, only fever, injection-site reactions, and allergic reactions seemed to be related to the vaccine. Events reported through the VAERS were similar in type and number for Vaqta and Havrix.

**Cost-effectiveness**

The cost-effectiveness of nationwide routine hepatitis A immunization has been evaluated. Compared with no childhood immunization against hepatitis A, routine immunization at 1 year of age would result in 183 806 fewer infections and 32 fewer deaths in each cohort.63 The cost-effectiveness ratio was estimated at $173 000 per life-year gained and $24 000 per quality-adjusted life-year (QALY) gained. When out-of-cohort herd immunity was considered, immunization at 1 year of age yielded a societal cost of $1000 per QALY gained. Another economic analysis that included the estimated re-
duction in secondary cases among household contacts of infected children yielded similar results. When these values are placed in context, the projected costs of implementation of a universal hepatitis A vaccine program is equivalent to an acellular pertussis vaccine program in adolescents and approximately 10% of the cost of the meningococcal vaccine program based on QALYs.

**Vaccine Administration and Storage**

Before administration, vaccine preparations should be shaken and, when well mixed, will be a slightly opaque, white-colored suspension. The vaccine should be administered intramuscularly with needle length based on age and size of the patient (see Table 1.5: Site and Needle Length by Age for Intramuscular Administration in Red Book). Hepatitis A vaccine in children is administered in a 2-dose series, with the first dose of Vaqta or Havrix administered to children as young as 12 months. The second dose of Havrix should be given 6 to 12 months after the first dose, and the second dose of Vaqta can be administered 6 to 18 months after the first dose. Twinrix is a 3-dose series given on a 0-, 1-, and 6-month schedule. If the immunization schedule for vaccines containing hepatitis A is interrupted, only the required immunization needs to be administered rather than restarting the series.

Hepatitis A vaccine is to be stored and shipped between 2 and 8°C (36 and 46°F). However, neither the immunogenicity nor reactogenicity of either Vaqta or Havrix was affected by storage at up to 37°C (98°F) for up to 1 week. The vaccine should not be frozen, because it will destroy vaccine potency.

**RECOMMENDATIONS**

As the next step in the incremental immunization strategy to prevent hepatitis A, the following recommendations are made.

**Children**

1. All children who live in the United States should receive hepatitis A vaccine at 1 year of age (ie, 12 to 23 months of age) as a 2-dose regimen. Immunization should be integrated into the routine childhood immunization schedule and completed according to the approved schedules (Table 1) using Havrix or Vaqta hepatitis A vaccines. Administration of 2 doses of the same hepatitis A vaccine is preferable. However, data indicate that the vaccines are interchangeable; thus, the 2-dose series may be completed with either of the vaccine preparations approved for children.

2. States, counties, and communities with existing hepatitis A immunization programs for children 2 to 18 years of age are encouraged to maintain these programs and expand to include children who are 12 to 23 months of age. In these areas, new efforts focused on routine immunization of preschool children should enhance, not supplant or replace, ongoing programs that are directed at a broader population of children.

3. In areas without existing hepatitis A immunization programs, catch-up immunization of unimmunized children 2 to 18 years of age can be considered. Such programs might especially be warranted in the context of increasing incidence or ongoing outbreaks among children or adolescents.

4. Immunocompromising conditions are not a contraindication to receiving hepatitis A vaccine. The preparation is an inactivated virus and has not been shown to result in any increased safety risks when administered to people with primary or secondary immunodeficiencies.

5. The vaccine should not be administered to people with a hypersensitivity to any of the vaccine components such as aluminum hydroxide and phenoxethanol.

**Persons at Increased Risk of Hepatitis A Virus Infection**

1. Children not previously immunized against hepatitis A virus who will be traveling to or living in areas with intermediate or high endemicity for the infection should be immunized before departure. Areas for which hepatitis A immunization is recommended before travel can be found at www.cdc.gov/travel/vaccinat.htm. Protection is reliably present by 4 weeks after administration of the first dose of hepatitis A vaccine and may afford protection as soon as 2 weeks after immunization.

2. Both adolescent and adult males who have sex with men should be immunized against hepatitis A virus.
Preimmunization serologic testing is not recommended for adolescents or young adults.

3. Immunization is recommended for users of either injectable or noninjectable illicit drugs. Again, preimmunization serologic testing is not recommended for adolescents or young adults.

4. Although changes in clotting-factor–preparation practices and donor screening have greatly reduced the risk of acquiring hepatitis A for recipients of clotting factors, susceptible individuals should be immunized against hepatitis A before administration of the clotting factors.

5. Susceptible persons who work with hepatitis A virus in a laboratory setting should be immunized against the virus.

REPORTING ADVERSE EVENTS
The safety of hepatitis A vaccines will continue to be assessed through ongoing monitoring of data from the VAERS and other surveillance systems. Any adverse event suspected to be associated with hepatitis A immunization should be reported to the VAERS. Information on how to report adverse events is available at www.fda.gov/cber/vaers/vaers.htm. VAERS forms also can be obtained by telephone at 800-822-7967.

FUTURE NEEDS AND RESEARCH
Ongoing Disease Surveillance
The incidence of hepatitis A is at an all-time low in the United States.5 The decrease in the rate of disease is temporally associated with implementation of an immunization program against the infection. Although it is likely that immunization has been a major contributor to the decrease, other factors, including improved hygiene or cycling of disease, which have been characteristic of the epidemiology of hepatitis A virus in the past, could have contributed to the decreased hepatitis A incidence during the past decade. Comprehensive information on hepatitis A immunization coverage is vital to fully evaluate the effect that immunization has had on reduction of the incidence of the infection. Unfortunately, only limited data are available, and immunization coverage among adults is not assessed systematically.

Although data from the Third National Health and Nutrition Examination Survey69 contributed to the understanding of hepatitis A, the information gained focused on prevalence of infection. As vaccine implementation increases, it will be most important to collect prospective data on disease incidence by fully investigating disease outbreaks and encouraging national reporting of cases through state and local health departments to the National Notifiable Diseases Surveillance System. Only with the availability of these data will it be possible to determine the full impact and added value of these mass-immunization programs.

Potential Need for a Catch-up Schedule
Strategies for catch-up immunization are often components of universal immunization programs. However, currently available data from the United States, Israel, and some European countries suggest that there is a significant herd-immunity effect associated with immunization of young children against hepatitis A.6,7,70 A mandate for catch-up immunization should await further surveillance to determine if the indirect effect on older, nonimmunized groups continues.

Observing for Need of a Booster Dose for Adults
Hepatitis A infection in childhood typically is mild, and in children younger than 6 years, 90% of infections are asymptomatic. However, acquisition of the infection during adolescence and adulthood typically is associated with symptomatic infections that can be debilitating for weeks. Thus, for a hepatitis A immunization program to be effective, the vaccine has to confer long-term protection. Otherwise, an asymptomatic childhood infection could be replaced by symptomatic disease after exposure later in life. Because the vaccine has been commercially available for only 10 years, data on the persistence of antibody is based largely on information collected from trials that tested the immunogenicity and efficacy of the vaccine. However, the available data are encouraging, finding that protective antibody concentrations persist more than 10 years after immunization.71 In addition, mathematical modeling suggests that protective antibody concentrations may persist for more than 25 years after immunization.59 Finally, studies suggest that immunity may be present even beyond the ability to detect circulating antibodies.71 Thus, although it is not considered necessary for booster immunization in a fully immunized healthy person, observation will be needed to determine if this evidence-based recommendation will need to change over time.

Evaluation of Immunogenicity With Coadministration of Vaccines
Limited data are available that indicate that coadministration of hepatitis A vaccine with other vaccines in the recommended childhood series does not affect immunogenicity of the vaccines.54 Additional studies to assess hepatitis A vaccine immunogenicity in conjunction with other vaccines, particularly varicella and pneumococcal conjugate vaccines, will need to be collected and evaluated. However, unless data become available to the contrary, one should assume that concomitant administration of hepatitis A vaccine with other recommended vaccines of children is safe and immunogenic.
Evaluating Vaccine Acceptance
The success of a hepatitis A vaccine program will depend on the enthusiasm that physicians and members of the health care team display toward implementation of the program. In addition, the attitudes of families to incorporation of “another vaccine” into the crowded immunization schedule will critically affect the program. In a 2003 survey, only 51% of age-eligible children who lived in the 11 states with recommendation to be immunized with hepatitis A vaccine had received at least 1 dose of vaccine.72 Only 1% of children who lived in states without a hepatitis A vaccine recommendation had received 1 dose of vaccine. Despite this moderate uptake of vaccine in “vaccine” states, there was a significant decrease in hepatitis A disease, suggesting the vaccine is highly effective and, along with the excellent safety profile of the vaccine, can provide the encouraging information needed to sustain a vaccine program.

ACKNOWLEDGMENTS
This American Academy of Pediatrics statement was prepared in parallel with the recommendations of the CDC report “Prevention of Hepatitis A Through Active or Passive Immunization: Recommendations of the Advisory Committee on Immunization Practices (ACIP)” (MMWR Recomm Rep. 2006;55[RR-7]:1–23). Much of the background presented in this statement is based on literature review, analysis of unpublished data, and deliberations of the CDC staff in collaboration with the Advisory Committee on Immunization Practices Hepatitis Working Group. We acknowledge the CDC Division of Viral Hepatitis and especially Drs Beth Bell and Tracy Liu of the Hepatitis Working Group for outstanding technical expertise and review.

REFERENCES
27. Mutsch M, Spicher VM, Gut C, Steffen R. Hepatitis A virus
32. Koff RS. Clinical manifestations and diagnosis of hepatitis A virus infection. Vaccine. 1992;10(suppl 1):S15–S17
41. Tedeschi MV, Yoshida CF, Silva M, Feinstone SM. Detection of hepatitis A viral genome in stool samples of patients with relapsed hepatitis A by the polymerase chain reaction. Mem Inst Oswaldo Cruz. 1989;84:429
42. Schiff ER. Atypical clinical manifestations of hepatitis A. Vaccine. 1992;10(suppl 1):S18–S20
44. Cooksley WG. What did we learn from the Shanghai hepatitis A epidemic? J Viral Hepat. 2000;7(suppl 1):1–3
68. Conner BA, Phair J, Sack D, et al. Randomized, double-blind study in healthy adults to assess the boosting effect of Vaqta or
"FOR JIMMY AND THE BOYS AND GIRLS OF AMERICA": PUBLICIZING CHILDHOOD CANCERS IN TWENTIETH-CENTURY AMERICA

“On the evening of 22 May 1948, Ralph Edwards, host of the popular radio program Truth or Consequences, introduced his audience to a special guest: ‘Tonight we take you to a little fellow named Jimmy. We’re not going to give you his last name, because he’s just like thousands of other young fellows and girls in private homes and hospitals all over the country.’ Without further explanation, the program commenced as Edwards prompted Jimmy to list his favorite Boston Braves players. Members of the team’s starting lineup filed into his hospital room one by one, and presented the boy with autographed baseball memorabilia. Jimmy then joined the men in singing ‘Take Me Out to the Ballgame’ on air and received special permission to attend a game the next day—a day designated as ‘Jimmy’s Day’ at the ballpark. After his young guest signed off, Edwards told listeners that Jimmy was a twelve-year-old undergoing cancer treatment in Boston. He asked them to contribute money toward a television set for the boy’s room and, more generally, to aid ‘Jimmy and the boys and girls of America.’ Members of the show’s audience responded generously, reportedly donating more than $200,000 to the fund and sending tens of thousands of ‘get well’ cards to Jimmy. By drawing upon child-centered fund-raising strategies pioneered by other earlier voluntary health agencies, the Jimmy Fund and its mission to direct research and treatment toward childhood cancers were launched with overwhelming public support.”

Noted by James W. Kendig, MD
Preparation for Emergencies in the Offices of Pediatricians and Pediatric Primary Care Providers

Committee on Pediatric Emergency Medicine

ABSTRACT

High-quality pediatric emergency care can be provided only through the collaborative efforts of many health care professionals and child advocates working together throughout a continuum of care that extends from prevention and the medical home to prehospital care, to emergency department stabilization, to critical care and rehabilitation, and finally to a return to care in the medical home. At times, the office of the pediatric primary care provider will serve as the entry site into the emergency care system, which comprises out-of-hospital emergency medical services personnel, emergency department nurses and physicians, and other emergency and critical care providers. Recognizing the important role of pediatric primary care providers in the emergency care system for children and understanding the capabilities and limitations of that system are essential if pediatric primary care providers are to offer the best chance at intact survival for every child who is brought to the office with an emergency. Optimizing pediatric primary care provider office readiness for emergencies requires consideration of the unique aspects of each office practice, the types of patients and emergencies that might be seen, the resources on site, and the resources of the larger emergency care system of which the pediatric primary care provider’s office is a part. Parent education regarding prevention, recognition, and response to emergencies, patient triage, early recognition and stabilization of pediatric emergencies in the office, and timely transfer to an appropriate facility for definitive care are important responsibilities of every pediatric primary care provider. In addition, pediatric primary care providers can collaborate with out-of-hospital and hospital-based providers and advocate for the best-quality emergency care for their patients.

INTRODUCTION

Pediatricians and pediatric primary care providers (PPCPs) are vitally important members of the emergency care system for children. Children with potentially life-threatening illnesses and injuries are sometimes taken to primary care offices, which often serve as the child’s medical home, by parents or caregivers seeking help from health care professionals they know and trust. The office site then serves as an entry into the emergency care system, and it is there that vital, perhaps even life-saving, care is provided.

Studies have shown that emergencies are common in primary care practices that provide care to children. In 1 study, the authors surveyed 52 pediatric offices and found that these practices saw a median of 24 emergencies per year.1 Most of the offices (82%) reported that they encountered, on average, at least 1 emer-
gery per month. In another study, 62% of pediatricians and family physicians in an urban setting who were asked about emergencies in their offices reported that they assessed more than 1 patient each week in their offices who required hospitalization or urgent stabilization.2

Despite these findings, which suggest that a significant number of children present to primary care offices with urgent or emergent problems, some health care professionals discount the need for preparation because “emergencies are not very common” or because they feel they can rely on rapid response from emergency medical services (EMS) or proximity to a hospital. Some PPCPs have interpreted risk-management guidelines to mean that having emergency equipment and medications on site will increase their liability in emergency situations; however, lack of preparation may be a true cause of increased liability. Other providers state that emergency equipment and medications are expensive, and they cannot afford to maintain these items. Indeed, studies have shown that a substantial number of practices are not prepared to manage pediatric emergencies and have documented deficiencies in equipment and training.1,4 One study showed that physicians with training in advanced pediatric life support (APLS) were more likely to have resuscitation equipment and to have conducted a mock code in their office.4 Other studies have supported training in basic life support (BLS) as well as advanced life support (ALS), as suggested by the American Academy of Pediatrics (AAP) policy statement published in December 2004.3 The statement suggested that pediatricians will improve the chance of survival of children who experience cardiac arrest by advocating for cardiopulmonary resuscitation (CPR) training of parents and caregivers and participating in BLS training courses as participants and instructors.

STATEMENT OF THE PROBLEM
Although pediatric emergencies may not be common occurrences in all primary care settings, numerous studies have shown that children continue to be taken to primary care offices at the time of an emergency.6-9 The most common types of emergencies include respiratory emergencies, seizures, infections in young infants, and dehydration.10 Pediatricians and PPCPs may be required to provide urgent or emergent care in their offices for children with these conditions, at least until the arrival of EMS. The consequences of being unprepared are serious; therefore, appropriate stabilization of pediatric emergencies and timely transfer to an appropriate facility for definitive care are important responsibilities of every PPCP.11

OFFICE-BASED SELF-ASSESSMENT
Optimizing PPCP office readiness for emergencies begins with a consideration of the unique aspects of each office practice, the types of patients and emergencies that have been or might be seen, the resources on site, and the resources of the larger emergency care system of which the PPCP’s office is part. Reviewing a standardized office-based self-assessment can provide PPCPs with a starting point for optimizing office readiness.12 Sample questions include:

1. What emergencies have you experienced in the office setting? How often have office emergencies occurred in your practice?
2. What is your office setting (freestanding office, clinic based, health center based, hospital based, other)? Are there resources outside of your office that you could call on during an office emergency (eg, security, other medical or dental professionals in the same building, hospital code team)?
3. What are the high and low staffing points during the times when your office is open? (Include nights and weekends if applicable.) What is the emergency readiness training of the staff present during those times?
4. How far is your office from a site of definitive care, such as the nearest emergency department (ED) or the nearest pediatric center? How long does it take for EMS to respond to a 9-1-1 call from your office? What is the point of entry for your local 9-1-1 response team (ie, the facility to which they are required by protocol to bring a pediatric patient)?
5. Does your practice care for any children who are technology dependent or have special health care needs? Do you have need for any additional equipment or expertise should a technology-dependent child have an emergency in your office?
6. What is your risk-management company’s policy regarding emergency preparedness of your office?

Answers to these and other questions (see Appendix 1) can help PPCPs examine their office practice within the context of the larger emergency care system and make informed choices to enhance the readiness of their office setting for anticipated emergencies.

PARENT AND PATIENT EDUCATION
Through effective parent and patient education and anticipatory guidance, some emergencies that present to the PPCP office could be prevented or directed more appropriately to an ED. PPCPs can improve the outcome of childhood emergencies by advocating CPR and first aid training of parents and caregivers and by educating them about how to prevent injuries, recognize an emergency, and respond appropriately in terms of first aid, CPR, accessing the private office or EMS, and choosing the appropriate facility: office, urgent care center, local ED, or pediatric specialty care center. Anticipatory guidance regarding emergencies should include when and
how to access EMS (9-1-1 or the local emergency access number), posting the national Poison Control Center number (800-222-1222), a means of obtaining after-hours advice, the need for consent for treatment of minors, any constraints to emergency care from health plan requirements for referral, and what facilities to access in a true emergency. Family teaching materials such as The Injury Prevention Program, the first aid chart, and EMS information card are available through the AAP.15

PPCPs should discuss advance directives and limitation of life-sustaining treatment with a family before any emergency develops.19 Because some states do not allow EMS personnel to recognize and respect pediatric advance directives, it is critical that any out-of-hospital do-not-resuscitate or “accept-natural-death” orders be discussed at the time of their issue with local EMS medical directors to ensure that EMS personnel, when called and asked to perform comfort measures instead of aggressive resuscitative measures, are acting within preapproved medical direction and remain free from liability.

In addition, PPCPs who care for children with special health care needs can help improve emergency care for these children by providing a brief but comprehensive summary of important information for hospital and prehospital providers. Nationally recognized forms, such as the emergency information form,15 and medical-alert jewelry can provide needed information during an emergency. Inquiring about the existence of a local Emergency Medical Services for Children–sponsored “child alert” program can further enhance the EMS response and care by strengthening the link with responding EMS personnel and decreasing the anxiety levels of parents, EMS personnel, and hospital staff. With the family’s consent, mechanisms to identify children with special needs in an emergency can be established and shared with local EMS providers.15

PREPARING THE OFFICE AND OFFICE PERSONNEL

At the time of a pediatric emergency, good resuscitation knowledge and skills are essential to provide high-quality care and ensure the best chances for intact survival for the child, but the outcome does not depend solely on the pediatrician or primary care physician. Successful stabilization requires an effective team, and members of the office staff need to be prepared; they need adequate knowledge, training, and resources to respond to an emergency.10 They also need an opportunity to practice; awareness of each member’s role on the team and an opportunity to rehearse these roles will lead to a more highly functioning, effective emergency team.

The first person to assess patients who arrive in the office may be the least clinically experienced employee: the secretary or receptionist. These employees should be able to recognize emergencies and know how to summon help. They can be taught about signs and symptoms that may signal an emergency in a child, such as labored breathing, cyanosis, audible stridor or wheezing, grunting or flaring, seizures, depressed mental status, or uncontrolled bleeding.16 Front-desk personnel or the office nurse might periodically check the waiting area, especially if the waiting time for an acute care visit is prolonged or the waiting area is not under direct visual supervision.

A clear response plan, including a plan for those times when the office is open but not fully staffed, is very helpful at the time of an emergency.17 Each member of the office staff can have a specific role in the overall management plan, including designation of the individual who will access the emergency response system. Personnel who fulfill this role should receive training specific to accessing EMS, and they should be knowledgeable about the capabilities and level of response provided by the local EMS agency. Office staff will need to provide information to the EMS dispatcher, including office address and location of the office within the building; the child’s age, condition, and vital signs; the transport destination; and need for an ALS unit if available.11 Office staff cue cards can be posted by the telephone to assist in accessing emergency help and providing appropriate information12 (Appendices 2A–2D).

The PPCP can preassign roles for the “resuscitation team,” and the team can then practice these roles by participating in office mock codes or simulated exercises on a regular basis. The PPCP can “run the code” and provide medical direction, but a contingency plan should be developed to guide staff if no physician is in the office at the time of the emergency. Pediatric care protocols adapted from EMS providers might help provide a basis for the development of individualized office-based protocols and scenarios for the top 5 to 10 emergency conditions. Tasks of the office team during an emergency include assisting and performing resuscitative measures, such as chest compressions, and recording or documenting the events of the resuscitation process and drawing up and administering medications and fluids. It may be helpful for PPCPs to assess the skill level and knowledge of new employees and clinical care providers who will likely have different levels of experience in handling pediatric emergencies. All PPCPs in practice should have a minimum of BLS training, and a more advanced level of training is essential if the office does not have rapid access to an ALS response unit. When the office is open, there should be someone in the office who can recognize an emergency situation, provide BLS, and activate the emergency response system. PPCPs can facilitate training in BLS and ALS by providing time for employees to take training courses offered in the community or local hospital or by collaborating with local EMS personnel who can offer training courses on site at the office. By working together in nonemergency situations, EMS providers and office staff can create an opportunity to improve communication and develop teamwork skills that will
facilitate the transfer of care at the time of an emergency (see Appendix 3). EMS staff may be able to identify logistic problems, such as ease in locating the office or accessing the examination room with a gurney, and clarify treatment and destination protocols in their region. Some PPCPs have also found it very helpful to review actual cases and invite local EMS providers to participate in simulated drills and to supplement certification or training with teaching specific to the most common problems seen in their offices.

**EMERGENCY EQUIPMENT AND MEDICATIONS**

Trained personnel must have rapid access to appropriate equipment and medications to use at the time of an emergency. All office staff members need to know where resuscitation equipment is located so that no time is wasted in finding it during an emergency. For those who practice in an office located in or near a hospital, basic airway equipment may be all that is needed. However, for practices and offices that have prolonged emergency response times, stabilization efforts may need to be maintained for up to 30 minutes before EMS arrives with their equipment and stabilization skills. In these offices, more equipment might be required to maintain an airway and to initiate treatment of shock.

Resuscitation equipment can be kept in an examination room designated as the resuscitation room, which is prestocked in an organized way, or it can be stocked and organized in a box, to be taken to the site of the resuscitation. A list of recommended equipment for office emergencies is provided in Table 1, and a list of recommended medications is provided in Table 2. Equipment and medications should be checked on a regular basis to ensure that all essential items are present, operating properly, and not expired.

Health care professionals, patients, and families have developed an increased awareness of issues related to patient safety since the release of the Institute of Medicine report on medical errors in 1999.16 Current safety literature suggests that pediatric patients are especially susceptible to medication error (dosing error) because of the need to calculate doses rather than using standardized dosing as in adult medicine.19,20 Over the past few years, a number of clinical tools have been developed to help decrease medication errors. One of the most familiar is the Broselow pediatric resuscitation tape, which is now available in many EDs and offices across the country.21 Studies have shown that the Broselow tape can help to reduce medication dosing (prescribing) error by providing precalculated doses.22 It allows prescribers to avoid the step of mathematical calculation, a frequent source of error in the medication process.21–22 However, some studies have described a potential increase in medical errors when using the Broselow tape because of its design and the fact that it is often used incorrectly.23–27

The Duke University Medical Center maintains a Web site (Duke Enhancing Pediatric Safety Web site; available at: www.dukehealth.org/deps) that was developed to provide education about the proper use of the Broselow tape. New resuscitation tools, which are currently being developed, will help pediatricians and pediatric care providers by providing suggested care protocols with recommended medications and precalculated doses.

**TABLE 1 Recommended Equipment for Pediatric Office Emergencies**

<table>
<thead>
<tr>
<th>Office Emergency Equipment and Supplies</th>
<th>Priority&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Airway management</td>
<td></td>
</tr>
<tr>
<td>Oxygen-delivery system</td>
<td>E</td>
</tr>
<tr>
<td>Bag-valve-mask (450 and 1000 mL)</td>
<td>E</td>
</tr>
<tr>
<td>Clear oxygen masks, breather and nonrebreather, with reservoirs (infant, child, adult)</td>
<td>E</td>
</tr>
<tr>
<td>Suction device, tonsil tip, bulb syringe</td>
<td>E</td>
</tr>
<tr>
<td>Nebulizer (or metered-dose inhaler with spacer/mask)</td>
<td>E</td>
</tr>
<tr>
<td>Oropharyngeal airways (sizes 00–5)</td>
<td>E</td>
</tr>
<tr>
<td>Pulse-oximeter</td>
<td>E</td>
</tr>
<tr>
<td>Nasopharyngeal airways (sizes 12–30F)</td>
<td>S</td>
</tr>
<tr>
<td>Magill forceps (pediatric, adult)</td>
<td>S</td>
</tr>
<tr>
<td>Suction catheters (sizes 5–16F) and Yankauer suction tip</td>
<td>S</td>
</tr>
<tr>
<td>Nasogastric tubes (sizes 6–14F)</td>
<td>S</td>
</tr>
<tr>
<td>Laryngoscope handle (pediatric, adult) with extra batteries, bulbs</td>
<td>S</td>
</tr>
<tr>
<td>Laryngoscope blades (0–2 straight and 2–3 curved)</td>
<td>S</td>
</tr>
<tr>
<td>Endotracheal tubes (uncuffed 2.5–5.5; cuffed 6.0–8.0)</td>
<td>S</td>
</tr>
<tr>
<td>Stylets (pediatric, adult)</td>
<td>S</td>
</tr>
<tr>
<td>Esophageal intubation detector or end-tidal carbon dioxide detector</td>
<td>S</td>
</tr>
<tr>
<td>Vascular access and fluid management</td>
<td></td>
</tr>
<tr>
<td>Butterfly needles (19–25 gauge)</td>
<td>S</td>
</tr>
<tr>
<td>Catheter-over-needle device (14–24 gauge)</td>
<td>S</td>
</tr>
<tr>
<td>Arm boards, tape, tourniquet</td>
<td>S</td>
</tr>
<tr>
<td>Intravenous needles (16 and 18 gauge)</td>
<td>S</td>
</tr>
<tr>
<td>Intravenous tubing, microdrip</td>
<td>S</td>
</tr>
<tr>
<td>Miscellaneous equipment and supplies</td>
<td></td>
</tr>
<tr>
<td>Color-coded tape or preprinted drug doses</td>
<td>E</td>
</tr>
<tr>
<td>Cardiac arrest board/backboard</td>
<td>E</td>
</tr>
<tr>
<td>Sphygmomanometer (infant, child, adult, thigh cuffs)</td>
<td>E</td>
</tr>
<tr>
<td>Splints, sterile dressings</td>
<td>E</td>
</tr>
<tr>
<td>Automated external defibrillator with pediatric capabilities</td>
<td>S</td>
</tr>
<tr>
<td>Spot glucose test</td>
<td>S</td>
</tr>
<tr>
<td>Stiff neck collars (small/large)</td>
<td>S</td>
</tr>
<tr>
<td>Heating source (overhead warmer/infrared lamp)</td>
<td>S</td>
</tr>
</tbody>
</table>

Note that some offices are located at a distance from EMS services. Providers in offices that are located more than 10 minutes away from the nearest EMS service need equipment that may not be required in the initial minutes of a resuscitation but will be required as the resuscitation effort extends past 10 minutes.

<sup>a</sup> E indicates essential, S, strongly suggested (essential if EMS response time is >10 minutes).

HEALTH CARE PROFESSIONAL SKILLS

In the setting of a pediatric emergency, PPCPs must be able to provide basic airway management and initiate treatment of shock. The skills required to perform these tasks successfully are usually acquired in training, but many PPCPs do not use them frequently, because the incidence of office emergencies is not high. Nonetheless, when an emergency occurs, the best chance for intact survival of the child is determined by adequate airway management. Therefore, providers need to keep their resuscitation skills and knowledge up to date. Pediatric advanced life support (PALS) and APLS courses provide up-to-date pediatric advanced care skills and knowledge to providers. When planning a mock code for office personnel, designate a recorder for each simulated exercise. After completing the exercise, critique not only the mock code itself but also the documentation of the event. In addition, keep records of mock codes held in the office with a note of “lessons learned” from each one. If there has been a recent change in office practice or equipment (ie, new forms used to document treatment), it may be helpful to include these as specific teaching points after completing the exercise, critique not only the mock code itself but also the documentation of the event. In addition, keep records of mock codes held in the office with a note of “lessons learned” from each one. If there has been a recent change in office practice or equipment (ie, new forms used to document treatment), it may be helpful to include these as specific teaching points after completing the exercise.

Another strategy used by some offices to improve “readiness” for an emergency is a scavenger hunt. This may be especially helpful for new staff or employees as part of their orientation to the office setting. Staff members are given a list of items (such as emergency equipment, medications, supplies, posted protocols for accessing EMS) and asked to find them within a defined period of time.

DOCUMENTATION

The most effective tool for risk management of office emergencies is documentation of efforts taken to improve office readiness, such as purchase and maintenance of equipment and medications; training provided; and policy and practice for patient education, patient triage, and office flow. Working toward the common goal of improved outcomes for office emergencies, pedi-

---

**TABLE 2** Office Emergency Drugs

<table>
<thead>
<tr>
<th>Priority*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drugs</td>
</tr>
<tr>
<td>Oxygen</td>
</tr>
<tr>
<td>Albuterol for inhalationb</td>
</tr>
<tr>
<td>Epinephrine (1:1000)</td>
</tr>
<tr>
<td>Activated charcoal</td>
</tr>
<tr>
<td>Antibiotics</td>
</tr>
<tr>
<td>Anticonvulsant agents (diazepam, lorazepam)</td>
</tr>
<tr>
<td>Corticosteroids (parenteral/oral)</td>
</tr>
<tr>
<td>Dextrose (25%)</td>
</tr>
<tr>
<td>Diphenhydramine (parenteral, 50 mg/mL)</td>
</tr>
<tr>
<td>Epinephrine (1:10,000)</td>
</tr>
<tr>
<td>Atropine sulfate (0.1 mg/mL)</td>
</tr>
<tr>
<td>Naloxone (0.4 mg/mL)</td>
</tr>
<tr>
<td>Sodium bicarbonate (4.2%)</td>
</tr>
</tbody>
</table>

**Fluids**

<table>
<thead>
<tr>
<th>Priority*</th>
</tr>
</thead>
<tbody>
<tr>
<td>E</td>
</tr>
<tr>
<td>S</td>
</tr>
</tbody>
</table>

- E indicates essential; S strongly suggested (essential if EMS response time is more than 10 minutes).
- b Metered-dose inhaler with spacer or mask may be substituted.


---

The reception area of the office, holding an “infant” (mannequin) and complaining that the infant will not wake up. The receptionist would then need to activate the emergency response system designed for the office. In some offices, this may mean calling aloud for help; in other offices, the receptionist may ring a bell or overhead-page someone for help. The nurse can be instructed to respond as he or she would in a real emergency, perhaps by taking the infant to a treatment room if one exists in the office or by calling for help and locating the emergency equipment box to bring to the examination room where the infant is taken. Clinical staff can then be asked to locate specific pieces of equipment they may need for the resuscitation. For example, they might be asked to locate the oxygen tank with appropriate tubing and demonstrate how to turn it on or locate the bag-valve-mask device (eg, Ambu bag) and demonstrate proper bagging technique. The physical act of locating and handling equipment such as the bag-valve-mask device is important for staff members to practice to be better prepared to perform these tasks when a true emergency occurs. Team members can then offer observations of their own and others’ performances, and specific action plans for improvement and problem solving can be developed. Action plans might address such topics as additional training needs, skills practice, equipment needs, and organizational issues. A sample of a mock-code evaluation form is shown in Appendices 4A and 4B, and sample scenarios for use in a mock code are shown in Appendix 5.

When planning a mock code for office personnel, designate a recorder for each simulated exercise. After completing the exercise, critique not only the mock code itself but also the documentation of the event. In addition, keep records of mock codes held in the office with a note of “lessons learned” from each one. If there has been a recent change in office practice or equipment (ie, new forms used to document treatment), it may be helpful to include these as specific teaching points after the simulated exercise.
Pediatric practices can collaborate with their risk-management agent to find ways to reduce risk while improving readiness. Documentation should also be included in office training and mock codes and, most importantly, during true resuscitation attempts.

Emergency situations are the most difficult to document properly. Stress levels are high, there are often not enough trained assistants, and other patients in the waiting room cannot be ignored. However, complete and accurate information regarding resuscitative efforts is vital for ongoing care, especially at the time of transfer of care. Documentation should include the date and time of treatment, the estimated or actual weight of the child if known, medications given with dosages and response noted, fluid volumes given, information or explanations given to the family, and the condition of the child at the time of departure from the office. An example of a “resuscitation log” is shown in Appendix 6.

EMERGENCY MEDICAL SERVICES

When a child requires resuscitation in an office, the PPCP and office staff members need help from other members of the emergency care team to ensure the best possible outcome. EMS personnel can provide competent assistance to the office team.

EMS personnel who respond to pediatric emergencies may include first responders, BLS emergency medical technicians (EMTs), or ALS EMTs (eg, EMT-paramedics). First responders and BLS EMTs can offer essential BLS skills and transport. ALS EMTs, acting under medical control and advanced protocols, can perform advanced airway-management skills, including positive-pressure ventilation and placing airway adjuncts. They can also establish intravenous or intraosseous access, administer intravenous or nebulized medications, defibrillate, and perform other advanced skills, in accordance with local protocols. Because only a small percentage (5%-10%) of EMS calls are for pediatric patients, many paramedics may have limited experience in working with children. PPCPs can help EMS personnel gain experience with children by inviting them to observe well-child visits in the office and providing an opportunity to interact with children. In many communities, paramedics have assisted pediatricians by helping to teach PALS or CPR classes to office staff. Establishing good and close communication with local EMS providers can help inform your office of their unique skill sets and introduce them to the types of emergencies to which they might be called to respond from your office. Including the medical director of the EMS service in office-based emergency-preparedness activities can assist in helping the EMS personnel be prepared with proper training and protocols for pediatric patients.

EMS personnel are well trained in resuscitative skills and are important members of the health care team. However, they cannot assist in the care of children who are critically ill unless they are called. PPCPs should confirm the access number for EMS (usually 9-1-1, but in some areas it may still be a 7-digit number) and have the number posted for easy access by any office staff directed to call EMS when an emergency is recognized. The office staff and physician should not delay activating EMS because of a concern that they might not actually be needed. In the long run, it is much better to have a unit respond even if the call is canceled en route or the child is not transported if he or she stabilizes in the office.

ADVOCACY

PPCPs have a critical role as advocates for high-quality emergency care for their pediatric patients. In partnership with out-of-hospital and hospital-based staff, PPCPs can help ensure the readiness of all components of the emergency care system to care for children. For example, PPCPs can collaborate with local EMS to offer life-support training courses; provide office-based pediatric training for EMTs; participate in development of pediatric protocols with EMS; serve as advisors for out-of-hospital pediatric care review; and advocate for EMS to obtain appropriate pediatric training, equipment, and supplies. Finally, they can work to educate parents and lawmakers about the unique needs of children and the special and sometimes complex medical needs of children within the EMS system.

SUMMARY

Pediatricians and other PPCPs are critically important members of the pediatric emergency care team. They can be most successful when they understand their role within a larger emergency care system. Effective parent education can reduce emergencies and help ensure appropriate access to the emergency care system. Careful self-assessment of office practice and policies can optimize office readiness before an emergency. When the primary care office becomes the entry point into the EMS system for a child, that child’s long-term outcome can be greatly influenced by care provided in the first minutes of the emergency. Skilled physicians who work with appropriate equipment and a well-trained team, in collaboration with the EMS system, can achieve timely resuscitation and transfer to definitive care and offer the best chance for intact survival for every child and family who seeks their care in an emergency.

RECOMMENDATIONS

1. Perform a self-assessment of office readiness for emergencies based on a review of experiences of common emergent, urgent, and acute conditions treated in the office, including events involving children with special health care needs.
2. Develop an organizational plan for emergency response in the office, which includes:
   a. recognition of an emergency;
b. staff communication, roles, and responsibilities at the time of an emergency during times of high and low staffing;

c. protocol to access EMS; and

d. maintaining readiness through practice (mock codes).

3. Maintain recommended emergency equipment.

   a. Organize emergency equipment in a way that facilitates access to appropriate type and size at the time of an emergency.
   
   b. Develop a system to check equipment on a regular basis to make sure that it is immediately available and functioning properly.

4. Maintain recommended emergency medications and use a resuscitation aid or tool that provides suggested protocols with precalculated medication doses.

   a. Develop a system to check medications on a regular basis to make sure that stock is always present and expired medications are disposed of properly.

5. Develop a plan to provide education and continuing medical education for all staff.

   a. Front-line staff: recognizing emergencies; activating the emergency response plan; and understanding EMS roles, capabilities, and access
   
   b. Clinical staff: maintaining knowledge and skills related to pediatric emergencies
   
   c. All staff: maintaining readiness; participating in mock codes; office checklist; office self-assessment

6. Practice mock codes in the office on a regular basis (quarterly or biannually).

   a. Involve as many staff members as possible.
   
   b. Include documentation as a defined role for a staff member.
   
   c. Critique the simulation and maintain a list of lessons learned.
   
   d. Include EMS when possible.
   
   e. Include disaster-preparedness scenarios in mock drills (see www.dukehealth.org/deps).

7. Educate families about what to do in an emergency.

   a. Encourage first aid and CPR training for parents and caregivers.
   
   b. Provide access number for after-hours advice, emergency response system, and poison information to families.
   
   c. Educate families about symptoms and situations for which they should access office advice, EMS, and poison information.

   d. Facilitate use and maintenance of emergency information forms for children with special health care needs.

8. Partner with EMS and hospital-based emergency providers to ensure optimal emergency care and emergency/disaster readiness for children.

APPENDIX 1: SELF-ASSESSMENT OF OFFICE PREPAREDNESS FOR PEDIATRIC EMERGENCIES

As you answer these questions, you may be better able to identify those areas in which your office preparedness can be enhanced.

1. What emergencies have you experienced in the office setting? How often have office emergencies occurred in your practice?

2. What is your office setting (freestanding office, clinic based, health center based, hospital based, other)? Are there resources outside your office on which you could call during an office emergency (e.g., security, other medical or dental professionals in the same building, hospital code team)?

3. What are the high and low staffing points during the times when your office is open? (Include nights and weekends if applicable.)

4. What is the emergency readiness of the staff present during those times? (Include first aid, CPR, BLS, ALS, PALS, APLS, Emergency Nurse Pediatric Course, other continuing medical education, etc.)

5. Have nonclinical staff been trained to recognize a potential or actual emergency?

6. What anticipatory guidance and education do you provide parents regarding injury prevention, first aid and CPR training, recognizing and responding to emergencies, and accessing EMS?

7. Is your waiting room under direct observation or screened frequently by a clinical staff member? Is it childproofed?

8. Does your practice have a written protocol for response in an office emergency? Does that protocol cover times of low staffing?

9. Do all staff members know how to access the EMS system? Staff members should be able to give the location and directions to the office, level of clinical staff present, age and condition of child (including vital signs if appropriate), desired transport location, and the level of emergency response (ALS or BLS) required.

10. Do you have specific telephone triage protocols for nonclinical and clinical staff?

11. How far is your office from a site of definitive care, such as the nearest ED, or the nearest pediatric center?
12. How long does it take for EMS to respond to a 9-1-1 call from your office?
13. Has EMS ever been to visit your office for a non-emergency call or to receive experience in evaluating pediatric patients?
14. What level of provider comes when you call 9-1-1: first responder, BLS, or ALS? Does your local EMS have the necessary equipment and expertise to manage children?
15. What is the point of entry for your local 9-1-1 response team (ie, the facility to which they are required by field protocol to bring a pediatric patient)?
16. If EMS does not go directly to a pediatric center on a 9-1-1 call, how do you emergently transport a child to the desired pediatric center when necessary?
17. Does your office use oxygen? If so, how is it supplied? Do all clinical staff members know how to operate the oxygen canister and know where the key is kept?
18. What emergency dosage strategy do you use in the office (code card, length-based tape, dosage book, no strategy)?
19. What airway equipment do you stock? Do all staff members know how to locate, choose, and use the appropriate size of equipment for any given child?
20. What equipment and supplies do you have on site to provide you and your staff with universal precautions?
21. Does your practice care for any children who are technology dependent or have special health care needs? Do you have need for any additional equipment or expertise if a technology-dependent child should have an emergency in your office?
22. Do you have written office protocols for common office emergencies such as respiratory distress, anaphylaxis, sepsis, dehydration, and supraventricular tachycardia?
23. How do you document events during an office emergency (assigned role, tape recorder, retrospective, other)?
24. How do you and your staff maintain skills and readiness? (Examples include attending nursery deliveries, moonlighting in urgent care or pediatric ED, being a PALS or APLS instructor, holding regular mock office codes and scavenger hunts for infrequently used equipment, providing expert review of pediatric runs for your local EMS, or other.)
25. How do you document parent education, staff training, protocols, and stocking for emergencies?
26. What is your risk-management company’s policy regarding emergency preparedness of your office?
27. Are there other aspects of your office practice that you think could be improved to achieve fewer office emergencies and better outcomes?

APPENDIX 2A: RECEPTION DESK EMERGENCY CARD

The following signs and symptoms may signal an emergency:

- Extremely labored breathing
- Blue or pale color (cyanosis)
- Noisy breathing (wheezing or stridor)
- Altered mental status
- Seizure
- Agitation (in the parent)
- Vomiting after a head injury
- Uncontrolled bleeding

If you feel a patient has symptoms that may signal an emergency, alert the following office staff: ________.

APPENDIX 2B: IMPORTANT TELEPHONE NUMBERS

- EMS provider (9-1-1 or your local emergency response number)
- Private ambulance service
- Specialized pediatric transport team(s)
- Office building security
- Police department
- Fire department
- Receiving hospital

Office address and directions ______

APPENDIX 2C: CALLING EMS FOR AN OFFICE EMERGENCY

Call 9-1-1 or your local EMS emergency response number: ________.

Be ready to give the emergency medical dispatcher the following information:

- Age and condition of child (with vital signs, if appropriate)
- Your office location (with directions and telephone number, if necessary)
- Level of clinical staff present
- Desired transport destination (pediatric center, local ED, other)
- Level of EMS provider required: ALS (advanced life support) or BLS (basic life support)
- If required, where security or other personnel will be meeting them to assist in guiding EMS to location of the child
APPENDIX 2D: IMPORTANT EMERGENCY TELEPHONE NUMBERS

(Fill in the blanks with your local emergency numbers)

<table>
<thead>
<tr>
<th>EMS</th>
<th>9-1-1 or local EMS access number</th>
</tr>
</thead>
</table>

Non-EMS Ambulance Transport Services


Pediatric Transport Teams


Referral Hospitals


Poison Control Centers


Helicopter Service


Police (non-9-1-1)


Security


Other


APPENDIX 3: BUILDING A PPCP-EMS
PARTNERSHIP—ACTION POINTS

- Offer your office as a pediatric training and refresher site for EMTs.
- Invite local EMS to participate in regularly scheduled office mock codes.
- Sponsor a local EMT to take a PALS instructor course together with one of your staff members.
- Consult your local EMS to review office emergency procedures, access, and equipment in light of their response time, medications, equipment, and destination options.
- Offer to review pediatric run sheets as part of your local EMS agency quality assurance/quality improvement processes.

APPENDIX 4A: MOCK-CODE EVALUATION FORM

<table>
<thead>
<tr>
<th>Clinical primary survey</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Airway assessed initially</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breathing then assessed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxygen started for respiratory distress</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Circulation assessed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial interventions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protocol or treatment guideline followed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient reassessed frequently</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary survey (head-to-toe examination)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All supplies requested were available</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supplies were found quickly when requested</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Code form” available and/or used</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personnel knew how to use equipment properly (O₂ tanks, etc)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protocols available and/or used</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leader communicated effectively</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Events recorded accurately</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roles were assigned</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Office staff reported to EMS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMS communicated needs/plans with office staff</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other comments</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX 5: SCENARIO SAMPLES

Diabetic ketoacidosis: 10-year-old with new-onset diabetic ketoacidosis; polyuria and polydipsia for 1 week; today lethargic and confused; glucose >800.

Sepsis: 2-year-old with meningococcemia; well in past but found this morning with rash, moaning and minimally responsive; had upper respiratory infection yesterday and 2 episodes of vomiting; otherwise fine.

Asthma: 8-year-old with asthma; has been wheezing for 2 days with upper respiratory infection but worsened this afternoon; told mom before he was brought to the office that he had been giving himself puffs of his inhaler every half hour most of the day.

Head trauma: 6-year-old with concussion and possibly more; was playing soccer and collided with another child; she was “out” for 2 to 3 minutes, then woke up and was groggy but oriented; vomited once on the way to your office.

Seizures: 1-year-old with a complex febrile seizure; pulling at her ears and found to have a temperature of 104°F; mom gave her a bath to cool her off, and she began to have a generalized seizure several minutes later; her parents rushed her to the office while carrying her on their laps; the seizure has persisted for over 20 minutes.

Stridor: 2-year-old with possible epiglottitis; woke up early this morning with very loud breathing and a barking cough; feels very hot to touch; has been drooling for past 30 minutes; now appears anxious and tired.

Anaphylaxis: 5-year-old boy who was stung by a bee while playing outside; mom notes that his eyes and lips swelled within minutes; she brought him to the doctor when he subsequently developed wheezing.
APPENDIX 6

Code chart: consider for use at time of office emergency. HR indicates heart rate; RR, respiration rate; BP, blood pressure; Pox, pulse oximetry; IO, intraosseous needle; IV, intravenous catheter.

---

AAP COMMITTEE ON PEDIATRIC EMERGENCY MEDICINE, 2005–2006

Steven Krug, MD, Chairperson
Thomas Bojko, MD, MS
Margaret A. Dolan, MD
*Karen Frush, MD
Patricia O’Malley, MD
Robert Sapien, MD
Kathy N. Shaw, MD, MSCE
Joan Shook, MD, MBA
Paul Sirbaugh, DO
Loren Yamamoto, MD, MPH, MBA

LIAISONS

Jane Ball, RN, DrPH
Susan Eads Role, JD, MSLS
    EMSC National Resource Center
Kathleen Brown, MD
    National Association of EMS Physicians
Kim Bullock, MD
    American Academy of Family Physicians
Dan Kavanaugh, MSW
Tina Turgel, BSN, RNC
    Maternal and Child Health Bureau
REFERENCES


ADDITIONAL REFERENCES

POLICY STATEMENT

School Transportation Safety

Committee on Injury, Violence, and Poison Prevention and Council on School Health

ABSTRACT

This policy statement replaces the previous version published in 1996. It provides new information, studies, regulations, and recommendations related to the safe transportation of children to and from school and school-related activities. Pediatricians can play an important role at the patient/family, community, state, and national levels as child advocates and consultants to schools and early education programs about transportation safety.

INTRODUCTION

School transportation plays a consistent and long-term role in the lives of children from preschool through high school. Pediatricians can participate by serving as resources, educators, consultants, and advocates for school transportation safety at the local, state, and federal levels. This revised policy statement provides updated recommendations that can enhance community systems for addressing safe transportation for children to and from school and school-related activities.

Expectations for school transportation and school bus safety should be upheld in an ongoing commitment from communities and states to ensure that children travel to and from school safely. The National Highway Traffic Safety Administration (NHTSA) School Bus Safety Program is committed to reducing school bus crashes, injuries, and fatalities. Congress has indicated that school transportation should be held to the highest level of safety. In addressing school transportation, all modes of travel must be considered, and measures must be taken to promote safety for each mode.

Modes of School Transportation

The Committee on School Transportation Safety of the Transportation Research Board studied the various modes of travel and associated risks for schoolchildren. Estimates of trips per year by mode of transportation during school hours were: passenger vehicle with adult driver, 45%; school buses, 25%, other buses, 2%; passenger vehicle with teen driver, 14%; bicycle, 2%; and walking, 12%. These estimates are limited, because they do not include school bus travel for extracurricular activities during or after normal school hours or during vacations. School bus crashes occur disproportionately on high-speed roads at night during transportation to and from extracurricular activities.

Annually during normal school travel hours, 23.5 million children are transported on 457,000 school buses, totaling 5.8 billion student trips and 3.13 billion miles. Each child who uses school bus transportation travels, on average, 1300 miles per school year. These estimates do not include school or school-related travel during nonschool hours.
School Transportation Injury

Annually, there are, on average, 815 student deaths and 152,250 injuries related to school travel during normal school travel hours. (These data are underestimates, because they do not include school-related trips or school bus crashes outside of school hours, and reporting is voluntary.) Two percent of the deaths and 4% of the injuries occurred in school buses. Seventy-five percent of the deaths and 84% of the injuries occurred in passenger vehicles. The fatality rates descend in the following order: (1) passenger vehicles with teen drivers, 55%; (2) passenger vehicles with adult drivers, 20%; (3) walking (pedestrians), 16%; (4) bicyclists, 6%; and (5) school and other buses, 2%. The injury rates descend in the following order: (1) passenger vehicles with teen drivers, 51%; (2) passenger vehicles with adult drivers, 33%; (3) walking (pedestrian), 6%; (4) bicyclists, 5%; and (5) school and other buses, 5%.3

The Fatality Analysis Reporting System includes fatality data on all school bus-related crashes, not just those during school hours. In the year 2001, 141 persons were killed. Of the fatalities, 16% were pedestrians, 9% were school bus passengers, 4% were school bus drivers, 3% were bicyclists, and the rest (68%) were occupants of other vehicles or other nonmotorists. Of the 22 child pedestrian fatalities, 82% were struck by the school bus. Data from the General Estimates System indicate that 13,000 persons are injured annually in school bus crashes.7 Of those injured, 46% (5980) were school bus occupants, 8% were school bus drivers, 38% were occupants of other vehicles, and fewer than 0.05% each were pedestrians, pedal cyclists, and nonmotorists.8 However, the National Transportation Safety Board (NTSB) determined that school bus crash data are incomplete and that injuries cannot be reliably estimated.9 The first emergency department–based study of nonfatal school bus–related injuries found that the number of injuries (17,000 annually to children 0–19 years of age) greatly exceeded previously published estimates. Motor vehicle crashes were the most frequent injury mechanism.10

RECOMMENDATIONS

School Bus Travel

The National Traffic and Motor Vehicle Safety Act of 1966 and the upgrades in the School Bus Safety Amendments of 1974 authorize the Department of Transportation to issue minimum standards for new school buses manufactured for sale in the United States.11 There are 35 federal motor vehicle safety standards (FMVSSs) that apply to school buses. Large school buses that carry more than 16 passengers are not required to be equipped with seat belts. The long-standing American Academy of Pediatrics (AAP) recommendation that seat belts be installed on all new school buses is further discussed later in this statement. Small school buses (weighing <10,000 pounds) built in accordance with FMVSSs are equipped with lap belts. Vehicles, including multipurpose vehicles that carry 11 or more persons that are sold or leased for transporting students to or from school or school-related events, are required to meet the FMVSS requirements applicable to school buses. States may prescribe additional regulations that apply to the use of any vehicle used to transport preprimary, primary, and secondary school students.12

The AAP recommends that all guidelines for safe transportation of all preschool- and school-aged children be applied during all school and school-related trips regardless of the hours of operation.

Preschool-Aged Children

Many school systems provide transportation for preschool-aged children. The NHTSA studies demonstrated that preschool-aged children were safest when properly transported in child safety restraint systems that meet FMVSSs 213 and 225.13 In January 2001, the Department of Health and Human Services issued transportation safety requirements for Head Start transportation. Within 5 years, transportation was limited to school buses or “allowable alternate vehicles.” This provision has since been extended to June 30, 2007. That is, on July 1, 2007, all Head Start children must be transported in a compliant vehicle, unless a waiver has been granted. All vehicles must be equipped with a communication system for emergencies, first aid kit, fire extinguisher, and seat belt cutter. Children weighing 50 pounds or less were required to travel in FMVSS 213–approved child restraints; this has since been updated to apply to children under the weight threshold of FMVSS 213 for approved child restraints (currently 65 pounds). As of January 2004, vehicles must be equipped to use child restraints. Retrofit of lap belts or child-restraint anchorages to properly secure the child safety restraint system to the school bus seat is allowed and must be reinforced according to the applicable FMVSS.13 The driver must have a commercial driver’s license and undergo criminal background checks. As of January 2004, all vehicles must have a bus monitor. Each Head Start agency is required to provide pedestrian-safety education for parents and children. An extension to January 18, 2006, for implementation of the requirement to provide car safety seats and bus monitors was allowed for Head Start programs that filed an application by April 1, 2004. A final rule was published on October 4, 2006, authorizing the Department of Health and Human Services to issue waivers from this requirement to Head Start grantees. The NHTSA has a curriculum for child passenger safety technicians, materials available regarding proper use of child safety restraint systems in school buses, and child passenger training materials for school bus drivers.

The AAP has recommended and advocated that school districts provide height- and weight-appropriate
car safety seats and restraint systems that meet FMVSSs for all preschool-aged children. These systems include booster seats for which a 3-point belt is available for installation. The AAP also supports the Head Start transportation safety requirements.

The AAP further recommends that school-based as well as non–school-based child programs follow guidelines for safe transportation. This includes all early education and child care programs and applies to car-pool transportation as well. The AAP Moving Kids Safely in Child Care program is the first national occupant-protection curriculum for child care providers and administrators; it provides detailed guidelines for safe transportation of all children.

School-Aged Children: Occupant Protection on School Buses

Compartmentalization has been the occupant-protection system for children in large school buses for more than 30 years and was the only available protection before child-restraint systems and seat belts were available for use in the school bus environment. Compartmentalization is provided by seats that are closely spaced with high, energy-absorbing seat backs. Data from real-world crashes comparing seat belt use versus compartmentalization only do not exist. However, recent studies have revealed that compartmentalization does not offer optimal protection and is not consistent with current technology and messages for children and families regarding the use of car safety seats and seat belts in all motor vehicles.

The NTSB, through a series of crash investigations, determined that compartmentalization as a method of occupant protection on school buses is incomplete. Compartmentalization does not provide protection during lateral (side) impacts with vehicles of large mass or in rollover collisions, because passengers do not always remain completely within the compartment. The NTSB recommended the development and implementation of a seat and restraint system that restrains passengers in the seating compartment. The NTSB also recommended the development of performance standards and requirements for school bus occupant-protection systems on newly manufactured school buses. The NTSB further recommended on-board recording devices to facilitate improved data collection in crashes. For optimal protection of all children, the AAP concurs with these recommendations.

The NHTSA conducted a study of school bus occupant protection in 2000 and determined that lap/shoulder belts on school buses performed best in dummy crash-testing compared with unbelted occupants, compartmentalization, and lap belts. Head-injury measurements were significantly lower with use of lap/shoulder belts than for use of compartmentalization or lap belts. In crash tests, the lap/shoulder belt restraint systems effectively kept the dummies in their seats.

The State of California Vehicle Code requires newly manufactured school buses to have a lap/shoulder belt restraint system, effective 2004 for small school buses and 2005 for large school buses. At the time of this publication, the states of Florida, Louisiana, New Jersey, and New York and many local school districts have passed school bus seat belt laws.

The AAP recommends that all children travel in age-appropriate, properly secured child-restraint systems when transported in all motor vehicles, including school buses, to ensure the safest ride possible. The AAP further recommends that all newly manufactured school buses be equipped with lap/shoulder restraint systems that can also accommodate car safety seats, booster seats, and harness systems. The AAP recognizes the added benefit of improved student behavior and consistent habits of restraint use when traveling in motor vehicles. Policies on seat belt use have been found to improve student behavior and reduce driver distraction. School districts must ensure the appropriate education of administrators, students, teachers, drivers, and parents in the use of occupant-protection devices.

School Bus Safety Features

The AAP recommends that all school buses, including private, parochial, and contractual, that are used for school and all school-related activity transportation be in compliance with all applicable federal regulations. Buses built before 1977 should be retired from use, because they are deficient in several significant safety standards.

Effective December 2, 1993, the FMVSSs were revised to require mirrors to improve driver visibility in front of and along both sides of school buses. In addition, districts should consider installing strobe lights for use during reduced-visibility conditions, an external loud-speaker system to enable the driver to communicate with children outside the bus, and loading and backing alarms or pulsating backup horns. School bus blind areas created by school bus bodies or mirrors are considerable. Electronic sensor systems are available but have not been evaluated adequately to determine their effectiveness. The AAP recommends that blind spots created by mirror systems and other vehicle-design aspects should be addressed by improved technology designed to decrease both crash and pedestrian injury risks because of limited visibility of a child by the bus driver.

The Children’s School Bus Exposure Study, prepared for the California Air Resources Board, found that diesel buses can have significantly higher on-board diesel-related pollutant concentrations than other vehicles because of intrusion of the bus’s own exhaust into the cabin. Increased exposure from commuting by school bus was estimated to increase a child’s lifetime cancer risk by approximately 4%, increase the incidence of lower respiratory symptoms by approximately 6%, and increase daily hospitalizations for asthma by approxi-
mately 1%. Several states and local governments have adopted airborne toxic control measures that limit school bus idling and idling at schools.\textsuperscript{25,26} Bus idling also contributes to poor indoor air quality inside schools from unfiltered air that enters through open doors. The AAP recommends that states adopt measures to protect school-aged children from exposure to toxic air contaminants.\textsuperscript{27} Additional measures to reduce children’s exposure to vehicle-related pollutants include replacement of older buses, use of alternate-fueled or particulate-trap-equipped buses, retrofitting buses with better emission-control technologies, minimizing bus caravanning, use of cleaner buses on longer routes, having passengers sit at the front of the bus if it is not full, and minimizing idling.\textsuperscript{24,28}

**School Bus Transportation of Children With Special Medical Needs**

Children with special needs and who are older than the preschool-aged child and require special restraint systems should be evaluated individually to determine the most appropriate system that meets their needs for positioning during travel, regardless of their age, weight, and height. Specific recommendations are outlined in the AAP policy statement on school bus transportation of children with special needs.\textsuperscript{29}

The use of wheelchairs is common for school bus transportation of children with disabilities. The AAP recommends that states adopt the requirements for use of wheelchairs on school buses outlined in the 1995 National Standards for School Buses\textsuperscript{32} and the AAP policy statement on school bus transportation of children with special needs.\textsuperscript{29}

**School Bus Driver Selection, Training, and Performance**

The Transportation Research Board stated that variations in school bus driver recruitment, selection, training practices, and rates of pay are likely to be associated with variations in driver safety performance.\textsuperscript{3} In another report, the Transportation Research Board recommended that all states provide formal training for school bus drivers, including training on school bus driver responsibility in ensuring safety of the children inside the bus and in loading zones.\textsuperscript{4}

The AAP believes that national standards for the selection, training, and regulation of school bus drivers should be established and implemented to ensure optimal driver performance.

To meet basic requirements, school bus drivers should:

- maintain a valid commercial driver’s license;
- be at least 21 years of age;
- show proof of an annual health history, assessment, and physical examination, including vision and hearing assessments, that document the absence of conditions that may compromise driving and child supervision;
- successfully complete a written or oral test covering driver duties, bus-operating procedures, traffic and school bus laws and regulations, record keeping, emergency and crash-related procedures, first aid, basic appreciation of the developmental stages and needs of preschool and school-aged children, child-supervision responsibilities, and transportation of passengers with special needs;
- maintain a satisfactory driving record as determined by the school district;
- successfully pass a review for a criminal record (including convictions of child sexual abuse and incidents or arrests for driving under the influence of alcohol or other drugs) that is reviewed annually; and
- pass a test for illicit drugs and alcohol as required by the district (mandatory testing is recommended if it is not already required).

To demonstrate operational and driving skills, school bus drivers should:

- pass a driving performance test and demonstrate safe loading and unloading procedures;
- demonstrate physical capability to successfully accomplish student evacuation; and
- demonstrate correct use of all occupant-protection systems that may be available on the school bus, including use of car safety seats, seat belt systems, and occupant-protection systems that are used by children with special medical or health needs.

Children with conditions such as anaphylactic allergies, severe asthma, diabetes, attention-deficit/hyperactivity disorder, autism or pervasive developmental disorder, and other chronic conditions may have health and safety issues during transport to and from school and school-related events. For that reason, the following are important:

- Drivers should be included in school plans for children with special medical and transportation needs.
- School bus drivers need to be aware of and prepared to intervene appropriately to ensure the safety of the individual child as well as all children on the trip. Interventions may require training beyond basic first aid.

**School Bus Passenger Instruction**

Passengers of all ages need to be taught safe riding and pedestrian behavior regardless of the frequency of school bus use. Instruction should include safe pedestrian practices going to and from the bus stop; safe behavior while waiting for the bus; safe practices for boarding and dis-
embarking the bus; safe behavior on the bus, including the use of child-restraint systems and seat belts when present; and procedures for emergency situations. Escort services for children crossing streets and roads should be considered.30

School Bus Passenger Supervision
Adult supervision on school buses should focus on ensuring that passengers stay seated and use age-appropriate car safety seats, seat belts, and other occupant-protection systems; ensuring that passengers keep their arms and heads inside the windows; assisting in emergency circumstances; assisting passengers with special needs; and escorting children across roadways. A second adult (other than the driver) serving as a monitor on the school bus can best meet these objectives. The Transportation Research Board states that it is generally agreed that monitors would enhance safety and reduce injuries by 25% to 75%; however, the cost estimate is high ($1.9 billion).4

School Bus Routes and Stops
Bus routes should be planned so that the bus does not have to back up, traffic disruptions are minimized, good fields of vision are provided at all stops, and the need for children to cross a street to board or leave the bus is minimized.4 Escorting children across streets has the greatest potential for injury reduction.4 Roads, traffic flow, traffic-control devices, and speed-limit enforcement should be maintained to optimize the safety of children.

Bicyclist and Pedestrian Travel to and From School
The motor, cognitive, and behavioral characteristics and abilities and limitations of children of different ages must be considered when assessing supervision needs necessary for students walking to and from school. There is no evidence that generic pedestrian-safety education is effective in reducing pedestrian injury. Bicyclists should be required to wear bicycle helmets properly.31 Children using nonmotorized vehicles for school and school-related trips should be required to use safety equipment, including helmets.32 Bicycle helmet use laws and enforcement increase helmet use.33 Driver education in school zones, including drivers who drop off and pick up students, must be addressed. Most drivers exceed speed limits in school zones.34 Safe Routes to School, an international movement, promotes infrastructure, environmental measures, enforcement, policy change, and education to enhance and promote safe walking.35

School-Zone Improvements
School-zone improvements would enhance the safety of all schoolchildren whether they walk, bike, take the school bus, or are dropped off and picked up with a passenger vehicle. These measures include marked drop-off and pick-up areas that are separate from school buses, school-zone speed-limit enforcement at 25 miles/hour, development of safe routes to school, and well-trained adult crossing guards. Crossing guards have been effective in improving pedestrian safety and have improved speed compliance and traffic control.36 The NHTSA issued guidelines for a uniform approach for traffic controls for school areas that were designed to enhance the safety of pedestrians. These guidelines further recommend that a school-route travel plan be developed systematically by school, law enforcement, and traffic officials.37 A multidisciplinary approach, involving school administrators, parent-teacher organizations, city planners, and law enforcement that includes infrastructure design as well as education of both students and drivers, offers potential to decrease death and injury to children in school zones.38

The AAP recommends the implementation of measures to improve the environmental infrastructure, including student supervision and crossing guards.

The Pediatrician’s Role
The pediatrician should promote school transportation safety at 4 levels: patient and family, community, state, and national. Pediatricians can serve as child advocates and consultants to child care and schools about transportation safety.

For school bus travel, the AAP emphasizes its long-standing position that seat belts be installed on all newly manufactured school buses. Three-point seat belts provide the best protection for school-aged children who have outgrown car safety seats.

Patient and Family Counseling
1. When addressing child passenger safety, inform families that the AAP has guidelines and policy statements for safe transportation of schoolchildren in school buses and other vehicles used for preschool, school, and child care transportation. In particular, inform parents that the AAP recommends that all children who travel in school buses use age-appropriate child-restraint systems and 3-point seat belts when they have outgrown child-restraint systems. Pediatricians should nevertheless counsel parents that large school buses, even when not equipped with seat belts, are the safest mode of school transportation.
2. Inform patients and families about the importance of bicycle helmets and other safety measures for children riding bicycles.
3. Inform parents that teens traveling together, especially with a teen driver, to and from school and to school-related events are at high risk of crash involvement and injury.
4. Promote passage and parent and community enforcement of graduated driver licensing laws, which
reduce fatal crash involvement of 16-year-old drivers by 16% to 21%.39,40

Community Role
5. Serve as a consultant to local parent groups, transportation directors, or school boards on the physical, cognitive, and psychosocial development of children as related to school transportation. Provide AAP guidelines and policy statements related to school transportation and teen driving.
6. Provide resources for communities to address safe routes for children who walk or bike to school.
7. Promote mandatory requirements for children to use bicycle helmets.
8. Advocate implementation of the recommendations of applicable policy statements at local school district meetings. Advocate for school districts to enforce graduated driver licensing laws.
9. Work with communities to plan for the transportation of children in planning new school sites and modifying existing sites.
10. Advocate for 3-point seat belt systems in all newly manufactured school buses

State Role
11. Serve as a consultant to state directors of school transportation to ensure that children’s needs and AAP guidelines are addressed in school transportation plans.
12. Advocate for mandatory bicycle helmet use laws and enforcement.
13. Share information from AAP policy statements.
14. Serve as a resource and consultant to the state department of education regarding training of bus drivers in areas relating to child passenger safety and child development and behavior.
15. Serve as a resource and consultant to the state department of education on pedestrian and bicycle safety for schoolchildren.

National Role
16. Encourage research to support continued improvement in school bus design and school-zone safety.
17. Advocate for mandated complete collection and reporting of data on fatalities and injuries by school districts and school bus transportation companies for all crash and noncrash events involving the school bus and multipurpose vehicles.

COMMITTEE ON INJURY, VIOLENCE, AND POISON PREVENTION, 2005–2006
Gary A. Smith, MD, DrPH, Chairperson
Carl R. Baum, MD
M. Denise Dowd, MD, MPH

Dennis R. Durbin, MD, MSCE
H. Garry Gardner, MD
Robert Sege, MD, PhD
Michael S. Turner, MD
Jeffrey C. Weiss, MD
Joseph L. Wright, MD, MPH

PAST COMMITTEE MEMBERS
Marilyn J. Bull, MD
Phyllis F. Agran, MD, MPH

liaisons
Ruth A. Brenner, MD, MPH
National Institute of Child Health and Human Development
Stephanie Bryn, MPH
Health Resources and Services Administration/Maternal and Child Health Bureau
Julie Gilchrist, MD
Centers for Disease Control and Prevention
Alexander “Sandy” Sinclair
National Highway Traffic Safety Administration
Jonathan D. Midgett, PhD
US Consumer Product Safety Commission
Lynne J. Warda, MD
Canadian Paediatric Society

STAFF
Rebecca Levin-Goodman, MPH

COUNCIL ON SCHOOL HEALTH, 2005–2006
Barbara Frankowski, MD, MPH, Chairperson
Rani S. Gereige, MD, MPH
Cynthia J. Mears, DO
Robert D. Murray, MD
Michele M. Roland, MD
Thomas L. Young, MD
Linda M. Grant, MD, MPH
Daniel Hyman, MD
Harold Magalnick, MD
George J. Monteverdi, MD
Evan G. Pattishall, III, MD

liaisons
Nancy LaCursia, PhD
American School Health Association
Donna Mazyck, MS, RN
National Association of School Nurses
Mary Vernon-Smiley, MD, MPH
Centers for Disease Control and Prevention
Robin Wallace, MD
Independent School Health Association

staff
Su Li, MPA

*Lead author
REFERENCES


17. State of California, Department of Motor Vehicles. Vehicle Code §27316


30. State of California, Department of Motor Vehicles. Vehicle Code 22112


38. State of California, Office of Traffic Safety and Safe Routes to


RESOURCE LIST

American Academy of Pediatrics, Healthy Child Care America. Moving Kids Safety in Child Care. Elk Grove Village, IL: American Academy of Pediatrics; 2002 [A curriculum for child care providers and administrators that covers seat restraints, route planning, air bag safety, car safety seats, booster seats, and more is provided.]


US Department of Transportation, Federal Highway Administration. Safe routes to school Web site. Available at: http://safety.fhwa.dot.gov/saferoutes [The goal of Safe Routes to School is to empower communities to get more children traveling safely to school on foot or bike with greater frequency. Available resources include, but are not limited to, information on federal funding, state contacts, and links to other Web sites including the National Safe Routes to School Clearinghouse.]

International Walk to School Web site. Available at: www.iwalktoschool.org

US Environmental Protection Agency. Clean school bus USA. Available at: www.epa.gov/otaq/schoolbus [provides clinicians and communities information on how to reduce pollution caused by school buses]
Prevention of Varicella: Recommendations for Use of Varicella Vaccines in Children, Including a Recommendation for a Routine 2-Dose Varicella Immunization Schedule

Committee on Infectious Diseases

ABSTRACT
National varicella immunization coverage using the current 1-dose immunization strategy has increased among vaccine-eligible children 19 through 35 months of age from 27% in 1997 to 88% by 2005. These high immunization rates have resulted in a 71% to 84% decrease in the reported number of varicella cases, an 88% decrease in varicella-related hospitalizations, a 59% decrease in varicella-related ambulatory care visits, and a 92% decrease in varicella-related deaths in 1- to 4-year-old children when compared with data from the prevaccine era. Despite this significant decrease, the number of reported cases of varicella has remained relatively constant during the past 5 to 6 years. Since vaccine effectiveness for prevention of disease of any severity has been 80% to 85%, a large number of cases of varicella continue to occur among people who already have received the vaccine (breakthrough varicella), and outbreaks of varicella have been reported among highly immunized populations of schoolchildren. The peak age-specific incidence has shifted from 3- to 6-year-old children in the prevaccine era to 9- to 11-year-old children in the postvaccine era for cases in both immunized and unimmunized children during these outbreaks. Outbreaks of varicella are likely to continue with the current 1-dose immunization strategy.

After administration of 2 doses of varicella vaccine in children, the immune response is markedly enhanced, with ≥99% of children achieving an antibody concentration (determined by glycoprotein enzyme-linked immunosorbent assay) of ≥5 U/mL (an approximate correlate of protection) and a marked increase in geometric mean antibody titers after the second vaccine dose. The estimated vaccine efficacy over a 10-year observation period of 2 doses for prevention of any varicella disease is 98% (compared with 94% for 1 dose), with 100% efficacy for prevention of severe disease. Recipients of 2 doses of varicella vaccine are 3.3-fold less likely to have breakthrough varicella, compared with those who are given 1 dose, during the first 10 years after immunization.

To achieve greater levels of immunity with fewer serosusceptible people, greater protection against breakthrough varicella disease, and reduction in the number of outbreaks that occur nationwide among school-aged populations, a 2-dose varicella immunization strategy is now recommended for children ≥12 months of age.
• Children 12 months through 12 years of age should receive two 0.5-mL doses of varicella vaccine administered subcutaneously, separated by at least 3 months; if the second dose inadvertently is administered between 28 days and 3 months after the first dose, the second dose does not need to be repeated. All children routinely should receive the first dose of varicella-containing vaccine at 12 to 15 months of age. The second dose of varicella-containing vaccine is recommended routinely when children are 4 to 6 years of age (ie, before a child enters kindergarten or first grade) but can be administered at an earlier age.

• People ≥13 years of age without evidence of immunity, as defined in the “Recommendations” section of this report, should receive two 0.5-mL doses of varicella vaccine separated by at least 28 days. For people who previously received only 1 dose of varicella vaccine, a second dose is necessary to provide evidence of immunity.

Both a monovalent varicella vaccine (Varivax [Merck & Co Inc, Whitehouse Station, NJ]) and a combination quadrivalent varicella-containing vaccine (ProQuad [Merck & Co Inc] or measles-mumps-rubella-varicella) are licensed by the Food and Drug Administration for use in the United States. Monovalent varicella vaccine is approved for use in children ≥12 months of age (and, therefore, adolescents and adults as well), and measles-mumps-rubella-varicella is approved only for children 12 months through 12 years of age. Neither varicella-containing vaccine contains thimerosal or other preservatives. When all vaccine components are indicated, combination vaccines are preferred whenever possible to minimize the number of injections.

BACKGROUND AND RATIONALE FOR RECOMMENDATIONS
Before licensure of the monovalent varicella vaccine (Varivax [Merck & Co Inc, Whitehouse Station, NJ]) in March 1995, approximately 4 million cases of varicella, 10 500 to 13 500 hospitalizations attributable to complications of varicella, and 100 to 150 deaths occurred annually.1–5 Beginning in 1996, routine varicella immunization was recommended for children between 12 and 18 months of age as well as for all susceptible people older than 19 months.6,7 In 1999, the Centers for Disease Control and Prevention (CDC) recommended a varicella immunization requirement for children attending child care and at elementary school entry,6 and in 2005 the CDC expanded this recommendation to include students from kindergarten through college.8 By July 2006, the District of Columbia and all states except Idaho, Montana, Vermont, and Wyoming had implemented a varicella immunization school-entry requirement that covers all or some of the recommended cohorts9 (see www.immunize.org/laws). Estimates of national varicella immunization coverage indicate an increase from 27% in 1997 to 88% by 2005 among vaccine-eligible children 19 to 35 months of age.10

These high immunization rates have had a dramatic effect on varicella disease burden, demonstrated by a 71% to 84% decrease in the reported number of varicella cases, an 88% decrease in varicella-related hospitalizations, a 59% decrease in varicella-related ambulatory care visits, and a 92% decrease in varicella-related deaths in 1- to 4-year-old children when comparing the prevaccine and postvaccine eras.5,11 Recent data suggest that in areas with vaccine coverage of approximately 90%, the incidence of varicella has declined 90% (CDC, written communication, 2005). However, vaccine effectiveness for prevention of varicella disease of any severity has been 80% to 85%11,12; as a result, large numbers of individual breakthrough varicella cases (defined as wild-type chickenpox occurring >42 days after immunization) continue to occur. In addition, outbreaks among highly immunized populations of schoolchildren continue to be reported.16–18 In these school outbreaks, varicella-vaccine coverage ranges from 96% to 100%, with vaccine effectiveness ranging from 72% to 85%. Immunized students with breakthrough varicella contributed to virus transmission, demonstrating that a 1-dose vaccine policy was not sufficient to control disease in these outbreaks. In addition, the peak age-specific incidence of varicella has shifted from 3- to 6-year-old children in the prevaccine era to 9- to 11-year-old children in the postvaccine era for both immunized and unimmunized children during these outbreaks.

Surveillance by the CDC and state health departments has demonstrated that the number of reported cases of varicella has been relatively constant during the past 5 to 6 years. Given this experience, it is likely that outbreaks will continue, given the effectiveness of a 1-dose immunization policy. As varicella immunization rates increase, most varicella cases occurring during outbreaks will be in people who have been immunized. However, this pattern does not indicate an increase in the rate of breakthrough disease or evidence of increasing vaccine failure.

When varicella develops in an immunized person (breakthrough varicella), the median number of skin lesions is generally <50, the duration of illness is shorter, and the incidence of fever is lower than that in an unimmunized person. Serious complications probably are reduced, although data are insufficient to establish whether a reduction in bacterial skin infections, pneumonia, or encephalitis is achieved. Approximately 25% of breakthrough cases result in an illness with ≥50 lesions, similar to that occurring in unimmunized children. The occurrence of breakthrough varicella raises a number of concerns:

• Approximately 15% of vaccine recipients (those with no or partial response to 1 dose of vaccine) remain at
increased risk of breakthrough disease. Despite local outbreaks, exposure to varicella-zoster virus (VZV) will become less common as vaccine uptake increases. Therefore, these susceptible children may be at risk of severe varicella associated with VZV infection in adolescence and adulthood.

- Although contagion may be reduced among children who experience breakthrough disease, results from carefully studied school outbreaks show that schoolchildren with breakthrough disease can serve as the index case for an outbreak. School outbreaks that involve immunized children continue to occur regularly in the United States, particularly among elementary school students, even in states with high coverage rates. Students who experience breakthrough disease are excluded from school for 3 to 5 days, parents may miss employment while caring for their sick children, and exposed susceptible schoolteachers who may be pregnant or students with contraindications to immunization must be identified and considered for postexposure prophylaxis. These outbreaks place an increased workload on state health departments.

- Because most varicella cases occur in highly immunized populations, there may be concern regarding vaccine efficacy and a misunderstanding by physicians or parents who may conclude that vaccine efficacy is declining. This misperception can lead to frustration among both parents and physicians, with erosion of confidence in the varicella-vaccine program, especially among people who perceive varicella as a mild illness of childhood.

- Because immunized children who experience breakthrough disease are coinfected with 2 VZV strains (wild and vaccine types), they may be at increased risk of zoster from the reactivated wild-type strain later in life, compared with vaccine recipients who do not experience breakthrough disease.

EPIDEMIOLOGY OF THE DISEASE

VZV is a highly contagious pathogen, with 80% to 90% of susceptible people exposed in a household setting developing clinical infection.19 Transmission occurs via direct contact, airborne droplets, or infected respiratory tract secretions. The virus initially enters a susceptible host through the upper respiratory tract or the conjunctivae. An infected host is contagious from 1 to 2 days before the onset of the rash until all skin lesions are crusted. Secondary cases in a family setting usually are more severe than the primary case, likely because of a higher viral inoculum resulting from a more intense exposure. VZV spreads less readily in tropical climates than in temperate climates.20 and as such, a higher proportion of adults from tropical countries are serosusceptible to VZV compared with adults from countries with cooler climates.

During the 1980s, 33% of varicella cases occurred in 1- to 4-year-old children, and 44% occurred in 5- to 9-year-old children;21 >90% of all cases of chickenpox occurred in children younger than 15 years of age. During the early and mid-1990s, the age-specific incidence of chickenpox shifted to younger ages, with 1- to 4-year-old children having the highest incidence of infection rather than 5- to 9-year-old children.22,23 Data from the third National Health and Nutrition Examination surveys seroprevalence study from 1988 to 1994 indicate that, in the immediate prevaccine era, 95.5% of people 20 to 29 years old, 98.9% of people 30 to 39 years old, and >99.6% of people ≥40 years had immunity to VZV.24

In the prevaccine era, 97% to 99% of adults with a positive history of varicella were seropositive, and the majority of adults with negative or uncertain histories were seropositive (range: 71%–93%).25–28 No published data were available on the predictive value of a positive history of varicella disease in children during that time period. History of varicella may be becoming less reliable in the vaccine era, with only 75% of unimmunized children 1 to 4 years of age who report a positive history of chickenpox actually being seropositive.29 A second episode of chickenpox in a person is uncommon and occurs more frequently in immunocompromised hosts.30

After primary varicella infection (chickenpox), the virus establishes latency in neuronal ganglia. Reactivation of latent VZV causes herpes zoster (shingles). Approximately 20% to 30% of people over a lifetime develop herpes zoster, with disease incidence increasing markedly beginning at approximately 50 years of age.31–33 This increase in incidence of herpes zoster is associated with a relative loss of cell-mediated immunity to VZV that occurs naturally with aging.34–36 Herpes zoster in children is rare, although children who acquire chickenpox during the first year of life have an increased risk of shingles.37 Herpes zoster occurs more frequently in immunocompromised patients.38 Available data indicate that the risk of herpes zoster after immunization seems to be lower than the risk of zoster after wild-type varicella infection.39

CLINICAL MANIFESTATIONS OF THE DISEASE

After an average incubation period of 14 to 16 days (range: 10–21 days), clinically apparent disease ensues, with characteristic skin lesions in varying stages of development and resolution. Cutaneous lesions begin as macules and rapidly progress to papules, vesicles, pustules, and scabs. Fever and rash last approximately 5 days, with rash being more concentrated on the trunk and head than on the extremities. With wild-type disease, most children develop 250 to 500 skin lesions, and lesions frequently develop in the mouth, conjunctivae, or other mucosal sites. Bullous or hemorrhagic lesions
occur rarely. Elevations in hepatic transaminase levels occur relatively commonly during the acute illness.

Varicella in pregnant women may result in VZV transmission to the fetus or newborn. Intrauterine VZV infection may result in congenital varicella syndrome or clinical varicella in the newborn. Congenital varicella syndrome may consist of low birth weight, cutaneous scarring, limb hypoplasia, microcephaly, chorioretinitis, and cataracts. It occurs at a rate of <1.5% among infants born to women who contract VZV in the first 28 weeks of gestation. Children infected by VZV in utero during the second half of gestation can experience inapparent varicella and then develop zoster early in life without having had extracutaneous chickenpox. The onset of varicella in pregnant women 5 days before to 2 days after delivery may result in severe varicella in 17% to 30% of their newborn infants, which, if untreated, has a high mortality rate.40

Bacterial superinfection of skin lesions and bacterial pneumonia are among the most frequent complications of varicella in immunocompetent hosts. The viral skin lesions and associated pruritus predispose the infected person to *Staphylococcus aureus* superinfection, including methicillin-resistant *S. aureus* (MRSA) infection. Varicella is an important risk factor for invasive group A streptococcal infections, including those that result in necrotizing fasciitis. The most common central nervous system complication of wild-type VZV infection is transient cerebellar ataxia, although encephalitis, viral meningitis, and transverse myelitis also can occur. Reye syndrome is associated with aspirin use in children with varicella and influenza but is rare now that aspirin is used infrequently in the pediatric population. Neutropenia and thrombocytopenia can occur 1 to 2 weeks after initial infection. In immunocompromised patients, disseminated varicella with severe and even fatal outcomes is possible; people with defects in cell-mediated immunity, such as organ transplant recipients and HIV-infected people, are at highest risk of complications.

**VACCINE**

**Description**

Two varicella-containing vaccines are licensed for use in children and adults in the United States for prevention of varicella, both of which contain the live-attenuated Oka strain of VZV.41 Varivax is a monovalent vaccine that was licensed by the US Food and Drug Administration (FDA) in March 1995 for use in people 12 months and older.52 ProQuad (MMRV; Merck & Co Inc) combines the varicella vaccine with attenuated measles-mumps-rubella (MMR) vaccine viruses and was licensed by the FDA in September 2005 for use in children 12 months through 12 years of age.43 In addition to these 2 products, Zostavax (Merck & Co Inc) is a higher-concentration Oka/Merck strain vaccine that was licensed by the FDA in May 2006 and reduces the risk of herpes zoster and postherpetic neuralgia in people 60 years and older.33,44,45

Monovalent varicella virus vaccine is lyophilized and should be stored frozen at an average temperature of −15°C or colder until it is reconstituted for injection. After reconstitution as directed in the package insert, it should be kept at room temperature for a maximum of 30 minutes. Reconstituted vaccine contains a minimum of 1350 plaque-forming units (PFU) of Oka/Merck VZV in each 0.5-mL dose. Each 0.5-mL dose also contains 12.5 mg of hydrolyzed gelatin, trace amounts of neomycin and fetal bovine serum, 25 mg of sucrose, and trace residual components of MRC-5 cells (including DNA and protein). More than 50 million doses have been distributed in the United States since licensure.

The MMR components of the combination MMRV vaccine are identical and of equal concentration to those in the trivalent M-M-R II vaccine (Merck & Co Inc). However, the amount of Oka/Merck VZV in MMRV is higher than in the monovalent varicella vaccine, with a minimum of 3.99 log_{10} PFU in each dose of MMRV, compared with a minimum of approximately 3.13 log_{10} PFU in each dose of monovalent varicella vaccine. The other constituents are similar between the 2 products. The lyophilized product should be stored frozen at an average temperature of −15°C or colder until it is reconstituted for injection. After reconstitution as directed in the package insert, it should be kept at room temperature for a maximum of 30 minutes. As with the monovalent vaccine, MMRV does not contain thimerosal or other preservatives. More than 2 million doses have been distributed in the United States since licensure.

**Immunogenicity**

Commercially available assays for detection of antibody to VZV include the enzyme immunoassay and latex agglutination test.46,47 Two sensitive assays, gpELISA (glycoprotein enzyme-linked immunosorbent assay) and FAMA (sensitive fluorescent antibody to membrane antigen), have been used in clinical studies but are not commercially available.48 Commercially available enzyme immunoassay and latex agglutination tests are less sensitive and, therefore, unreliable in detecting immunity among immunized people. In addition, the latex agglutination test can yield false-positive results.49

The concentration of varicella antibody as measured by gpELISA 6 weeks after immunization correlates with neutralizing antibody concentration, VZV-specific T-lymphocyte proliferative responses, and protection against breakthrough varicella after exposure to VZV.50–54 Among children who have varicella antibody titers of ≥5 gpELISA units per mL 6 weeks after immunization, the vaccine efficacy rate is 95.5%, compared with an efficacy rate of 83.5% in children with <5 gpELISA units per mL.51 Approximately 76% to 90% of children who are immunized with a single dose of varicella-containing
Humoral and cellular immunity are important for viral clearance and for protection against reinfection or reactivation of latent VZV. Studies using the FAMA assay indicate that a titer of >1:4 at 16 weeks after immunization also correlates with protection against disease; fewer than 1% of healthy persons achieving a titer of >1:4 develop varicella after a household VZV exposure, compared with an attack rate of 55% among persons whose titers are <1:4. After 1 dose of monovalent varicella vaccine, 76% of healthy children seroconvert, as measured by FAMA assay, suggesting that many breakthrough cases of varicella in children who have received 1 dose of varicella vaccine may be attributable to primary vaccine failure rather than waning immunity.

Administration of 2 doses of monovalent varicella vaccine 3 months apart initially produces a higher geometric mean titer, compared with 1 dose of vaccine, and a larger percentage of patients who receive 2 doses have initial antibody concentrations of ≥5 gpELISA units per mL (99.6% vs 85.7%). Recipients of a second varicella-vaccine dose have fewer cases of breakthrough varicella and increased vaccine efficacy. When a second dose of varicella vaccine is administered 4 to 6 years after the first dose, a rise in VZV-specific antibody concentrations occurs rapidly and overall VZV antibody concentrations achieved are comparable with antibody concentrations achieved when 2 vaccine doses are administered 3 months apart.

Varicella, measles, mumps, and rubella antibody concentrations after administration of a single dose of MMRV vaccine are comparable with concentrations after administration of MMR vaccine and monovalent varicella vaccine concomitantly at separate injection sites. Among 5446 healthy children 12 to 23 months of age enrolled in 4 clinical trials, 91.2% achieved varicella antibody titers of ≥5 gpELISA units per mL (95% confidence interval: 90.3%–92%). Among a subset of 1035 healthy children who received a second dose of MMRV approximately 3 months after the first dose, 99.4% achieved varicella antibody titers of ≥5 gpELISA units per mL (95% confidence interval: 98.7%–99.8%). The varicella geometric mean titers increased approximately 41-fold after the second dose of MMRV vaccine.

Seroconversion rates for adolescents (≥13 years of age) and adults range from 75% to 94% after 1 dose of varicella vaccine but increase to 94% to 99% after a second dose of vaccine administered 4 to 8 weeks after the first. No published data are available on the proportion of adolescents or adults who achieve a gpELISA concentration of ≥5 U/mL after either 1 or 2 doses of varicella vaccine. Persistence of antibody over time period. One study indicates that for adults who develop breakthrough disease after exposure to varicella, neither attack rates nor severity of illness increase over time.

Efficacy and Effectiveness

Before vaccine licensure, 1 placebo-controlled clinical trial that used an earlier formulation of the Oka/Merck vaccine than the currently licensed product was conducted in the United States. This vaccine was 98% efficacious in preventing varicella after 2 years of follow-up and was 92% efficacious after household exposures. In open-label prelicensure studies in which efficacy was calculated by using historical attack rates for comparison, the efficacy of 1 dose of the varicella vaccine that ultimately was licensed most commonly ranged from 70% to 90% against infection and 95% against severe disease. In general, postlicensure effectiveness studies have reported a similar range for prevention against infection, with a few studies yielding lower values. The vaccine is highly effective (≥97%) in preventing severe varicella in postlicensure evaluations.

Fewer studies have evaluated efficacy and effectiveness of 2 doses of varicella vaccine. When 2 doses are administered 3 months apart, the estimated vaccine efficacy over a 10-year observation period for prevention of any varicella disease has been reported as 98.3%, with 100% efficacy for prevention of severe disease. The efficacy of a 2-dose regimen is significantly higher than that of a 1-dose regimen (94.4%; P < .001). Recipients of 2 doses of varicella vaccine are 3.3-fold less likely to have breakthrough varicella, compared with recipients of 1 dose (2.2% vs 7.3%; P < .001) during the first 10 years after immunization. Breakthrough cases developed annually in 0.0% to 0.8% of recipients of 2 vaccine doses, compared with 0.2% to 2.3% of recipients of 1 vaccine dose.

Safety

Studies conducted before and after licensure have confirmed that the varicella vaccine is safe and generally well tolerated. Pain and redness at the injection site were the only adverse events that occurred more frequently among vaccine recipients than among placebo recipients. In a study that examined a 2-dose regimen of monovalent vaccine separated by 3 months, there were slightly more injection-site complaints after the second dose. After 1 dose, recipients of MMRV were more likely...
than were recipients of monovalent varicella vaccine and MMR vaccine given at separate injection sites to have fever (21.5% vs 14.9%, respectively) and a measles-like rash (3% vs 2.1%, respectively).\(^4\) Both fever and measles-like rash usually occurred within 5 to 12 days of immunization, were of short duration, and resolved without long-term sequelae. After the second dose, there were no differences in incidence of fever or rash among recipients of MMRV compared with recipients of simultaneous MMR and varicella vaccines.

The risk of transmission of vaccine virus from immunocompetent vaccine recipients in whom varicella-like rash develops after immunization is extremely low, having been documented in only 5 cases, all of which occurred after exposures in household and institutional settings.\(^5\)–\(^7\) No cases of transmission have occurred after immunization of health care professionals. Therefore, the benefits of immunizing susceptible health care professionals outweigh this negligible or extremely low potential risk. As a safeguard, institutions may wish to consider precautions for personnel in whom rash develops after immunization and for those immunized personnel who will have contact with susceptible people at high risk of serious complications.

Concomitant administration of MMRV with diphtheria-tetanus-acellular pertussis (DTaP), *Haemophilus influenzae* type b (Hib), and hepatitis B vaccines generates comparable seroconversion rates as does individual administration of these vaccines.\(^4\) Data are absent or limited for concomitant use of MMRV with inactivated poliovirus, pneumococcal conjugate, influenza, and hepatitis A vaccines. Simultaneous administration of most widely used live-antigen and inactivated-antigen vaccines has produced seroconversion rates and rates of adverse reactions similar to those observed when the vaccines are administered separately. Therefore, monovalent varicella vaccine and MMRV vaccine may be administered simultaneously with other vaccines recommended for children 12 to 15 months and 4 to 6 years of age.

**Cost/Benefit**

Using current estimates of morbidity, mortality, direct costs, and indirect costs, analyses have found both 1-dose and 2-dose varicella immunization programs to be cost-saving at the societal level (F. Zhou, PhD, CDC, written communication, 2006). The incremental cost for the second dose is $96,000 per quality-adjusted life-year saved. When benefits from prevention of group A streptococcus infections and herpes zoster among immunized people are added to the model, incremental costs per quality-adjusted life-year saved are $91,000 and $17,000, respectively.

**Vaccine Storage and Administration**

Both monovalent varicella and MMRV vaccines must be stored frozen at an average temperature of \(-15^°C\ (5^°F)\) or colder. The diluent should be stored separately at room temperature or in a refrigerator. Both vaccines should be reconstituted according to directions in their respective package inserts. Both vaccines must be used within 30 minutes after reconstitution.

Monovalent varicella vaccine and MMR vaccine should be administered subcutaneously.

**RECOMMENDATIONS**

**Children 12 Months Through 12 Years of Age**

Both monovalent varicella vaccine and MMRV vaccine have been licensed for use in healthy children 12 months through 12 years of age. Children in this age group should receive two 0.5-mL doses of varicella vaccine administered subcutaneously, separated by at least 3 months (evidence grade I [see Appendix 1]). The recommendation for at least a 3-month interval between doses is based on the design of the studies that evaluated 2 doses in this age group\(^5\); if the second dose is administered inadvertently between 28 days and 3 months after the first dose, the second dose does not need to be repeated (evidence grade III). The American Academy of Pediatrics recommends the use of combination vaccines when all vaccine components are indicated to minimize the number of injections children receive.\(^7\)

All children routinely should receive the first dose of varicella-containing vaccine at 12 to 15 months of age (evidence grade I). The varicella vaccine should be administered to all children in this age range unless there is evidence of immunity to VZV (see “Documentation of Immunity”) or a contraindication to administration of the vaccine (see “Contraindications”). The second dose of varicella-containing vaccine is recommended routinely when children are 4 to 6 years of age (ie, before a child enters kindergarten or first grade) but can be administered at an earlier age (evidence grade III). A routine health maintenance visit at 11 to 12 years of age is recommended for all adolescents to evaluate immunization status and administer necessary vaccines, including varicella vaccine.

**People ≥13 Years of Age**

People ≥13 years of age without evidence of immunity should receive two 0.5-mL doses of varicella vaccine separated by at least 28 days (evidence grade I). The recommendation for at least a 28-day interval between doses is based on the design of the studies that evaluated 2 doses in this age group. For people who previously received only 1 dose of varicella vaccine, a second dose is necessary to provide evidence of immunity. Monovalent varicella vaccine, but not MMRV vaccine, is licensed for use in this age group.
Documentation of Immunity

Only doses of vaccine for which written documentation of the date of administration is presented should be considered valid. Neither a self-reported dose nor a history of immunization of the child as provided by a parent is, by itself, considered adequate documentation of immunity. A health care professional should supply an immunization record for a patient if that health care professional has administered the vaccine or has seen a record that documents immunization. People who lack either adequate documentation of immunization or other evidence of immunity should be immunized.

Although parental self-reporting of varicella disease has historically been considered valid enough to count as evidence of immunity, recent data on self-reporting in the varicella-vaccine era have revealed it to be less reliable than in the prevaccine era, probably because of the decrease in disease incidence and the proportion of mild cases among vaccine recipients, which are less readily recognized.

Serologic screening for VZV immunity generally is neither necessary nor recommended if a person has other acceptable evidence of immunity to the disease. With the exception of women who are known to be pregnant (see “Prenatal Screening and Postpartum Immunization”), people who lack acceptable evidence of immunity generally should be immunized without serologic testing. Postimmunization serologic testing to verify an immune response to varicella vaccine is not recommended, because available commercial assays are not sensitive enough and may give false-negative results.

Evidence of immunity to VZV in the pediatric population includes any of the following:

1. Documentation of 2 appropriately timed doses of varicella vaccine (evidence grade I)
2. Laboratory evidence of immunity or laboratory confirmation of disease (evidence grade I)
3. Varicella diagnosed by a health care professional or verification of history of varicella disease (evidence grade III)
   - For people reporting or presenting with typical disease, verification can be performed by any health care professional (eg, school or occupational clinic nurse, nurse practitioner, physician assistant, physician).
   - For people reporting or presenting with atypical and/or mild cases, assessment by a physician or physician’s designee is recommended, and 1 of the following should be sought: (a) an epidemiologic link to a typical varicella case or to a laboratory-confirmed case, or (b) evidence of laboratory confirmation, if it was performed at the time of acute disease. When such documentation is lacking, people should not be considered as having a valid history of disease, because other diseases may mimic mild atypical varicella.
4. History of herpes zoster diagnosed by a health care professional (evidence grade II-2)

Prenatal Screening and Postpartum Immunization

Prenatal screening of pregnant adolescent women for VZV immunity using the aforementioned criteria is recommended (evidence grade III). Varicella vaccine should not be administered to pregnant women, because the possible effects on fetal development are unknown, although no pattern of malformation has been identified after inadvertent immunization of pregnant women. After completion or termination of pregnancy, women who do not have evidence of VZV immunity should be immunized with the monovalent varicella vaccine before discharge from the hospital, birthing center, or abortion clinic; the second dose should be administered at least 28 days later (evidence grade III). Women should be counseled to avoid conception for 1 month after each dose of varicella vaccine.

A pregnant mother or other household member is not a contraindication for immunization of a child in the household (evidence grade III). Monovalent varicella vaccine should be administered to nursing mothers who lack evidence of immunity (evidence grade III).

Immunization of Immunocompromised Populations

General Recommendations

Varicella vaccine should not be administered routinely to children who have congenital or acquired T-lymphocyte immunodeficiency, including people with leukemia, lymphoma, and other malignant neoplasms affecting the bone marrow or lymphatic systems, as well as children receiving long-term immunosuppressive therapy. Certain children infected with HIV are an exception, as discussed later. Children with impaired humoral immunity may be immunized. Immunodeficiency should be excluded before immunization in children with a family history of hereditary immunodeficiency. The presence of an immunodeficient or HIV-seropositive family member does not contraindicate vaccine use in other family members.

When immunizing people with altered immunity against chickenpox (see “HIV Infection”), only monovalent varicella vaccine should be used. The Oka vaccine strain remains susceptible to acyclovir, and if a high-risk patient develops vaccine-related varicella, then acyclovir should be used as treatment.

Acute Lymphocytic Leukemia

Before routine immunization of healthy children against varicella was instituted in the United States in 1995, many young children with leukemia were susceptible to chickenpox. To protect them against serious and fatal
varicella, a research protocol for immunization against chickenpox was put in place, but this has since been terminated.\textsuperscript{79} Considering the variability of chemotherapy regimens and the current decreasing incidence of varicella in the United States, however, these high-risk children should not be routinely immunized. Immunization of varicella-susceptible leukemic children in remission should be undertaken only with expert guidance and with the availability of antiviral therapy, should complications occur.

Live-virus vaccines usually are withheld for an interval of at least 3 months after immunosuppressive cancer chemotherapy has been discontinued.\textsuperscript{80,81} The interval until immune reconstitution varies with the intensity and type of immunosuppressive therapy, radiation therapy, underlying disease, and other factors. Therefore, it often is not possible to make a definitive recommendation for an interval after cessation of immunosuppressive therapy when live-virus vaccines can be administered safely and effectively.\textsuperscript{81}

**HIV Infection**

Screening for HIV infection is not indicated before routine VZV immunization. After weighing potential risks and benefits, varicella vaccine should be considered for HIV-infected children in CDC class 1 with a CD4\textsuperscript{+} T-lymphocyte percentage of $\geq 15\%$ (evidence grade II-1).\textsuperscript{79} Eligible children should receive 2 doses of monovalent varicella vaccine with a 3-month interval between doses and return for evaluation if they experience a postimmunization varicella-like rash. With increased use of varicella vaccine and the resulting decrease in incidence of varicella in the community, exposure of immunocompromised hosts to VZV will decrease. As the risk of exposure decreases and more data are generated on the use of varicella vaccine in high-risk populations, the risk versus benefit of VZV immunization in HIV-infected children will need to be reassessed.

**Children Who Receive Corticosteroids**

Varicella vaccine should not be administered to people who are receiving high doses of systemic corticosteroids ($\geq 2$ mg/kg per day of prednisone or its equivalent or 20 mg/day of prednisone or its equivalent) for $\geq 14$ days (evidence grade III). The recommended interval between discontinuation of corticosteroid therapy and immunization with varicella vaccine is at least 1 month. Varicella vaccine may be administered to people on inhaled, nasal, and topical steroids.

**Households With Potential Contact With Immunocompromised People**

Transmission of vaccine-strain VZV from healthy people has been documented in 5 instances, resulting in 6 secondary cases. Even in families with immunocompromised people, including people with HIV infection, no precautions are needed after immunization of healthy children in whom a rash does not develop. Immunized people in whom a rash develops should avoid direct contact with immunocompromised susceptible hosts for the duration of the rash.

**Contraindications**

As generally with all vaccines, administration of varicella-containing vaccines is contraindicated in people with a history of severe (anaphylactic) reaction to the vaccine or its components (ie, neomycin or gelatin). Use of varicella-containing vaccines also is contraindicated in pregnant women and in people with known altered immunity (eg, HIV, hematologic and solid tumors, congenital immunodeficiency, and long-term immunosuppressive therapy) except as discussed previously. People with active untreated tuberculosis should not receive MMRV vaccine.

Of note, the monovalent varicella vaccine does not contain preservatives or egg protein, and although the measles and mumps vaccines included in MMRV are produced in chicken-embryo culture, the amounts of egg cross-reacting proteins are not significant. Therefore, children with egg allergy routinely may be given MMRV vaccine without previous skin testing.

**Precautions**

As with other vaccines, varicella-containing vaccines should not be administered to people who have moderate or severe illness, with or without fever. Recent receipt of blood products or immune globulin also is a precaution for administration of varicella-containing vaccines, as is a family history of immunodeficiency. Thrombocytopenia or a history of thrombocytopenic purpura are precautions for receipt of MMRV vaccine. For a detailed discussion of precautions, see the section on precautions and contraindications in the varicella chapter of the Red Book.\textsuperscript{79}

**Options for Postexposure Prophylaxis**

Depending on a person’s risk for serious varicella disease, options for postexposure prophylaxis include active immunoprophylaxis with a varicella-containing vaccine, passive immunoprophylaxis with VariZIG (the current formulation of varicella-zoster immune globulin, available under an investigational new drug application only) or immune globulin intravenous, or chemoprophylaxis with oral acyclovir. For a full consideration of these options, please refer to the Red Book.\textsuperscript{79}

**REPORTING ADVERSE EVENTS**

Clinically significant adverse events, regardless of whether they are suspected to have been caused by varicella-containing vaccine, should be reported to the Vaccine Adverse Event Reporting System. Forms can be obtained and submitted electronically through a secure
Web site (http://vaers.hhs.gov) or obtained by telephone at 800-822-7967.

**FUTURE NEEDS AND RESEARCH**

As 2-dose immunization schedules are introduced into clinical care, diligent monitoring of breakthrough disease and disease outbreaks will be critical. A 2-dose schedule is anticipated to substantially decrease disease outbreaks and breakthrough disease. If this is the effect, then wild-type VZV will circulate to an even lesser extent than it does now, and whether decreased circulation of VZV will contribute to waning immunity over time after receipt of 2 doses of a varicella-containing vaccine will need to be monitored.

**APPENDIX 1: US PREVENTIVE SERVICES TASK FORCE RATING SYSTEM OF QUALITY OF SCIENTIFIC EVIDENCE**

I. Evidence obtained from at least 1 properly designed, randomized, controlled trial.

II-1. Evidence obtained from well-designed controlled trials without randomization.

II-2. Evidence obtained from well-designed cohort or case-control analytic studies, preferentially from >1 center or group.

II-3. Evidence obtained from multiple time series with or without the intervention or dramatic results in uncontrolled experiments (such as the results of the introduction of penicillin treatment in the 1940s).

III. Opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees.

**COMMITTEE ON INFECTIOUS DISEASES, 2006–2007**

Joseph A. Bocchini, Jr, MD, Chairperson
Robert S. Baltimore, MD
Henry H. Bernstein, DO
John S. Bradley, MD
Michael T. Brady, MD
Penelope H. Dennehy, MD
Margaret C. Fisher, MD
Robert W. Frenck, Jr, MD
David W. Kimberlin, MD
Sarah S. Long, MD
Julia A. McMillan, MD
Lorry G. Rubin, MD

**LIAISONS**

Richard D. Clover, MD
American Academy of Family Physicians
Marc A. Fischer, MD
Centers for Disease Control and Prevention

---


**REFERENCES**


11. Centers for Disease Control and Prevention. National, state, and urban area vaccination coverage among children aged


57. Michalik DE, LaRussa PS, Steinberg SP, Wright P, Edwards K, Gershon AA. Primary immune failure after one dose of varicella vaccine are likely a cause of breakthrough infections in healthy vaccinated children. Presented at: 44th Annual Meeting of the Infectious Diseases Society of America; October 12–15, 2006; Toronto, Ontario, Canada. Abstract 146
71. Izurieta HS, Strebel PM, Blake PA. Postlicensure effectiveness of varicella vaccine during an outbreak in a child care center. JAMA. 1997;278:1495–1499
Stressing About Posttraumatic Stress Disorder

To the Editor.—

We are writing to you about a recent article that appeared in *Pediatrics*.1 The authors of this very interesting article claimed that posttraumatic stress disorder (PTSD) “symptoms and cortisol levels at baseline are associated with changes in hippocampal volume over an ensuing 12- to 18-month interval” after a stressful life event. However, we believe that this strong claim is not supported by the data the authors presented.

The first problem is that the methods the authors used to measure hippocampal volume are of low precision. In all studies that use MRI to measure brain volume, experimental variance can be introduced during data acquisition (eg, patient motion, changes in scanner hardware or software, scanner field variation) and during data analysis (eg, partial-volume problems, voxel misclassification, manual delineation error).2 Among the 15 children who they evaluated, the authors acknowledged (in line 3 of “Results”) that there was 1 child who showed an increase in right hippocampal volume that amounted to a full SD, or $\Delta 11011 15\%$. This child was excluded from additional evaluation. However, we believe that this strong claim is not supported by the data the authors presented.

The second problem is that the sample of 15 children evaluated was far too small to provide meaningful results. We have characterized the precision of measurement of brain volume in a cohort of 52 healthy adults imaged twice.3 Our goal was to determine the sample size needed to provide adequate statistical power in studies of brain volume by MRI. We imaged volunteers at weeks 0 and 12 using the same scanners and the same methods at both time points. We assumed that there would be no change in brain volume in the absence of an intervening neurologic event. Sample sizes necessary to yield 80% statistical power to detect a 5% change in brain volume were calculated for several experimental designs. The percentage difference between the 2 sets of 19 measurements averaged just 0.18%, but the required sample sizes, nevertheless, were far larger than anticipated.

The results shown suggest that, for a longitudinal study of hippocampal volume in adult patients and controls, a minimum sample size of 58 to 66 subjects is required for 80% statistical power, even if hippocampal volume had changed by 5%, which it did not (average change reported in this study was $+2.9\%$). It could be argued that the authors
actually presented a longitudinal 1-group study, which would require a sample size of only 16 to 18 subjects, but we think that such a study design is grossly inappropriate in subjects who are likely to be undergoing volumetric brain growth.

The third problem is that no data from healthy controls were presented, so we have no way of knowing what “normal” looks like by the methods that the authors used. This is a particularly serious error of omission in the case of children (the average age of subjects in their study was 10.4 years [range: 8–14 years]), in whom brain volume is likely to be changing. Furthermore, it would be a mistake to assume that hippocampal volume must be increasing simply because brain volume is still increasing in children under 12 years of age. In the absence of control data, we cannot assume that the reported decrease in hippocampal volume was abnormal. We note that it could be argued that analysis in this article was entirely focused on whether changes in hippocampal volume are correlated with PTSD symptoms, with hyperarousal, and with cortisol levels. Among 37 adult trauma survivors who were imaged prospectively within 1 week of trauma, there was no significant difference between the subjects who did and did not develop PTSD. Given that at least 1 earlier study included both pediatric patients and controls and followed the patients for a longer time period, the weight of evidence suggests that PTSD is not associated with a change in hippocampal volume, at least among children.

A fourth problem is that there were several problematic uses of statistical tests. For example, if 2 sets of brain-volume measurements are available, and all volumes in 1 set are exactly fivefold larger than those in the other set, the Pearson correlation would be 1.00, whereas a concordance correlation would be much lower, showing that the 2 sets of measures are not interchangeable. Last, an outlier was arbitrarily deleted from analysis when the statistical test used was not sensitive to outliers. Once a decision has been made to use Spearman correlations, there is no need to exclude outliers; all scores are turned into ranks, so a large change in absolute value has little impact.

A final problem with this article is that, although it was not rigorously done, it confirms 1 of 2 competing theories about how PTSD affects the hippocampus in children. As such, it may be given inordinate weight, although it is a weak study. An earlier longitudinal study of children that involved 18 subjects (rather than the 15 reported here) concluded that hippocampal volume did not differ between PTSD patients and controls at baseline, at follow-up, or over time. An earlier study by Carrion et al also failed to find a difference in hippocampal volume between pediatric patients with PTSD and controls. Among 37 adult trauma survivors who were imaged prospectively within 1 week of trauma, there was no significant difference between the subjects who did and did not develop PTSD. Given that at least 1 earlier study included both pediatric patients and controls and followed the patients for a longer time period, the weight of evidence suggests that PTSD is not associated with a change in hippocampal volume, at least among children.

We agree with Carrion et al that longitudinal studies are inherently stronger than cross-sectional studies and that the effects of PTSD on hippocampal volume in children can only be understood in the context of a strong longitudinal study. However, we do not believe that an underpowered, uncontrolled, and overanalyzed study is a step toward that goal.

### TABLE 1

Minimum Required Sample Size (Patients + Controls) for at Least 80% Statistical Power, Under a Variety of Assumptions About Study Design

<table>
<thead>
<tr>
<th>Structure</th>
<th>Cross-sectional: 2 Groups</th>
<th>Longitudinal: 1 Group</th>
<th>2 Groups (Change)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Power</td>
<td>N</td>
</tr>
<tr>
<td>Left hippocampus</td>
<td>250</td>
<td>80.2</td>
<td>16</td>
</tr>
<tr>
<td>Right hippocampus</td>
<td>276</td>
<td>80.0</td>
<td>18</td>
</tr>
</tbody>
</table>

The estimated statistical power is shown for a sample size that was calculated to yield at least 80% power.

ACKNOWLEDGMENTS

Dr Steen was supported by the National Alliance for Research on Schizophrenia and Depression as a Hoffmann Trust Investigator.

REFERENCES

2. Steen RG, Mull C, McClure R, Hamer RM, Lieberman JA. Brain...


doi:10.1542/peds.2007-0867

In Reply.—

Drs Steen and Hamer raise 5 questions about our recent article in Pediatrics: (1) the reliability of the hippocampus measurement; (2) sample size; (3) lack of a control group; (4) the number and type of statistical tests; and (5) the theoretical implications of its consistency with 1 of 2 competing hypotheses. We appreciate their interest in our work and address each of these issues below.

Regarding the issue of hippocampal reliability, we had strong interrater reliability with a 0.95 intraclass correlation coefficient. Our laboratory has published on hippocampal measurement approaches to address the inherent challenges of measuring small volumes (see ref 2). Imaging studies can always be affected by software and hardware; however, our laboratory has maintained backward compatibility with previous scanning techniques, and we routinely examine the effects of hardware or major software upgrades on image quality and characteristics to understand if there are any changes. We use the same scanner hardware and identical pulse sequences and software processing techniques. For example, we exclude scans that are technically unacceptable because of motion or other artifacts.

The second issue addresses sample size. We agree that any study can be improved with a larger sample size. Recognizing the limitation, we called it a pilot study. Steen and Hamer, however, raise the issue of power. Their comments do not apply to our findings, because the effect size obtained was large enough to meet the criteria set for statistical significance in the study. Their comments, however, do provide an important suggestion for future work in this area, specifically, that larger sample sizes and power may be particularly important for studies that attempt to replicate these findings in the future. In our article we stressed that our conclusions were “limited by the sample size” and that “findings should be considered preliminary until replicated.”

We agree with Steen and Hamer that performing this study with controls would have strengthened the results. As we mentioned in our article, however, our goal was to study the pathogenesis of hippocampal reduction in this sample and not to compare how hippocampal volume may differ between clinical and nonclinical groups. We have reported that in terms of functional impairment, there is no significant difference between children with posttraumatic stress disorder (PTSD) and children with subthreshold posttraumatic reactions.3 Studying the natural continuum of symptoms in traumatized youth and how it may impact brain development have been some of our goals.

The fourth issue is the statistical approach. We carefully explained in the article why we chose the statistical approach we took. We disagree that intraclass or concordance correlations would have been more appropriate, because we were attempting to establish the association between cortisol and PTSD symptoms and hippocampal change; we were not trying to show that PTSD and cortisol are “invariant in terms of mean and SD.”

Steen and Hamer also mentioned potential error caused by the multiple correlations presented in our Table 2 (type 1 error). The multiple tests were not probes for a significant finding among a host of possible associations but, rather, were theoretically directed tests. At least 20 of the 24 correlations presented in Table 2 were provided for completeness or descriptive or exploratory purposes only. What we find remarkable, from a theoretical perspective, was that both cortisol levels and PTSD symptoms at time 1 were associated with change in hippocampal volumes. If only PTSD symptoms or only cortisol levels were associated with change in hippocampal volumes, a chance association would seem more likely; however, this was not the case.

We note again that we made it clear in the discussion section of our article that ours was preliminary evidence and that larger studies are needed to confirm these findings and clarify the role that additional factors may play in this association. However, given the difficulty in conducting and obtaining funding for such studies without any evidence, we feel that our study was an important contribution at this point in the development of the literature, because it points to the need for such larger studies.

On their last point, Steen and Hamer cite 2 studies: one by DeBellis et al4 and another by our group.5 We would like to clarify their comments on these studies because, as presented, they can be misconstrued. In the DeBellis et al article, there were only 9 subjects with PTSD; the addition of 9 controls made it an 18-subject study. Our article, which showed no hippocampal differences, was a cross-sectional study. As we mentioned
Drs Steen and Hamer raise 5 questions about our recent article in Pediatrics: 1) the reliability of the hippocampal measurement; 2) sample size; 3) lack of a control group; 4) the number and type of statistical tests; and 5) the theoretical implications of its consistency with 1 of 2 competing hypotheses. We appreciate their interest in our work and address each of these issues below.

Regarding the issue of hippocampal reliability, we had strong interrater reliability with a 0.95 intraclass correlation coefficient. Our laboratory has published on hippocampal measurement approaches to address the inherent challenges of measuring small volumes (see ref 2). Imaging studies can always be affected by software and hardware; however, our laboratory has maintained backward compatibility with previous scanning techniques, and we routinely examine the effects of hardware or major software upgrades on image quality and characteristics to understand if there are any changes. We use the same scanner hardware and identical pulse sequences and software processing techniques. For example, we exclude scans that are technically unacceptable because of motion or other artifacts.

The second issue addresses sample size. We agree that any study can be improved with a larger sample size. Recognizing the limitation, we called it a pilot study. Steen and Hamer, however, raise the issue of power. Their comments do not apply to our findings, because the effect size obtained was large enough to meet the criteria set for statistical significance in the study. Their comments, however, do provide an important suggestion for future work in this area, specifically, that larger sample sizes and power may be particularly important for studies that attempt to replicate these findings in the future. In our article we stressed that our conclusions were “limited by the sample size” and that “findings should be considered preliminary until replicated.”

We agree with Steen and Hamer that performing this study with controls would have strengthened the results. As we mentioned in our article, however, our goal was to study the pathogenesis of hippocampal reduction in this sample and not to compare how hippocampal volume may differ between clinical and nonclinical groups. We have reported that in terms of functional impairment, there is no significant difference between children with posttraumatic stress disorder (PTSD) and children with subthreshold posttraumatic reactions. 

Studying the natural continuum of symptoms in traumatized youth and how it may impact brain development have been some of our goals.

The fourth issue is the statistical approach. We carefully explained in the article why we chose the statistical approach we took. We disagree that intraclass or concordance correlations would have been more appropriate, because we were attempting to establish the association between cortisol and PTSD symptoms and hippocampal change; we were not trying to show that PTSD and cortisol are “invariant in terms of mean and SD.”

Steen and Hamer also mentioned potential error caused by the multiple correlations presented in our Table 2 (type 1 error). The multiple tests were not probes for a significant finding among a host of possible associations but, rather, were theoretically directed tests. At least 20 of the 24 correlations presented in Table 2 were provided for completeness or descriptive or exploratory purposes only. What we find remarkable, from a theoretical perspective, was that both cortisol levels and PTSD symptoms at time 1 were associated with change in hippocampal volumes. If only PTSD symptoms or only cortisol levels were associated with change in hippocampal volumes, a chance association would seem more likely; however, this was not the case.

We note again that we made it clear in the discussion section of our article that ours was preliminary evidence and that larger studies are needed to confirm these findings and clarify the role that additional factors may play in this association. However, given the difficulty in conducting and obtaining funding for such studies without any evidence, we feel that our study was an important contribution at this point in the development of the literature, because it points to the need for such larger studies.

On their last point, Steen and Hamer cite 2 studies: one by DeBellis et al4 and another by our group.5 We would like to clarify their comments on these studies because, as presented, they can be misconstrued. In the DeBellis et al article, there were only 9 subjects with PTSD; the addition of 9 controls made it an 18-subject study. Our article, which showed no hippocampal differences, was a cross-sectional study. As we mentioned
in the introduction of the article, neither we nor DeBellis et al have found hippocampal differences when conducting cross-sectional studies. Two pediatric longitudinal studies have found either no hippocampal difference or a hippocampal reduction associated with cortisol and PTSD symptoms. The 2 approaches used by these 2 studies provided interesting leads into the study of hippocampal development and PTSD. Like any other subject in science, the ultimate answer to these questions relies on replication.

We hope our findings prompt large-scale investigations and serve as a basis for future exploration in this area.

Victor G. Carrion, MD
Department of Psychiatry and Behavioral Sciences
Stanford University
Stanford, CA 94305

Carl F. Weems, PhD
Department of Psychology
University of New Orleans
New Orleans, LA 70148

Allan L. Reiss, MD
Department of Psychiatry and Behavioral Sciences
Stanford University
Stanford, CA 94305

REFERENCES

doi:10.1542/peds.2007-1024

Increased Cerebral Blood Flow Velocity in Children With Sickle Cell Disease: Adenotonsillectomy or Transfusion Regimens?

*To the Editor.—*

We read with great interest the report from Hill et al., who showed evidence of a significant increase of cerebral blood flow velocity (CBFV) in children recruited from adenotonsillectomy waiting lists because of a history of snoring compared with controls. These features had major implications in the management of children with sickle cell disease (SCD) by underlining a relationship between abnormal cerebral high velocities and upper airway obstruction. Since the Stroke Prevention in Sickle Cell Anemia (STOP) trial, it has been recommended that children aged 2 to 16 years with SCD be screened by transcranial Doppler (TCD) to identify those at high risk for stroke who present with abnormal high velocities.2 These children at high risk are then offered transfusion regimens for stroke prevention.

We report here 2 boys with SCD, aged 6 and 5 years, whose abnormal, high velocities of \( \geq 200 \text{ cm/second} \) (middle cerebral artery) returned to normal values of \(<170 \text{ cm/second} \) after they had undergone adenoidectomy and/or tonsillectomy. Both patients had a history of snoring with adenotonsillar hypertrophy. Their hemoglobin values were 8 and 8.5 g/dL, respectively, and their cerebral MRI and magnetic resonance angiography findings were normal. Adenotonsillectomy was performed after 1 exchange transfusion and adenoidectomy after 1 simple transfusion; none of the patients experienced postoperative complications. No additional transfusions were provided. TCD was repeated 2 and 28 months after surgery in one of the patients and 3 months afterward in the other one; normal CBFV was found for both children.

High velocities are related to cerebral artery stenosis and severe anemia as well as tissue hypoxia. Adenotonsillar hypertrophy results in obstructive sleep apnea syndrome and nocturnal hypoxia, which is cured in 75% to 100% of children by adenotonsillectomy.3 Thus, upper airway obstruction may lead to a reversible increase of CBFV and must be carefully ruled out in children with SCD who have abnormal TCD findings before including them in transfusion regimens.

Brigitte Bader-Meunier, MD
Martine François, MD
Suzanne Verlhac, MD
Monique Elmahle, MD
Ghislaine Ithier, MD
Malika Benkerrou, MD
Assistance Publique-Hôpitaux de Paris
Hôpital Robert Debré
Paris 75019, France

REFERENCES
We appreciate the possible implications for children with sickle cell anemia (SCA) in the response by Bader-Meunier et al to our article on increased cerebral blood flow velocity in children with SCA with mild sleep-disordered breathing. Although there seems to be a link between sleep-disordered breathing and changes to the cerebral vasculature, the mechanism has not been elucidated. Candidates include the mechanical effects of snoring, associated hypoxemia, and infection/inflammation. Tonsillectomy is sometimes performed for recurrent tonsillitis or relief of upper airway obstruction in children with SCA, although mean overnight oxyhemoglobin saturation may not improve, and there are few systematically collected data that have addressed a relationship with SCA complications such as transcranial Doppler (TCD) abnormality.

TCD testing has been undertaken at least once in 384 children with homozygous SCA in the London cohort, and adenotonsillectomy has been performed on clinical grounds. In 183 (47.7%) patients, TCD data are available on 2 occasions (mean age at follow-up: 11.2 years [SD: 4.5 years]). Of these, 40 (21.9%) had undergone adenotonsillectomy between the 2 TCD measurements (mean age at surgery: 7.7 years [SD: 3.9 years]), 28 had adenotonsillectomy before the first TCD (not included in this analysis), and 115 had not undergone adenotonsillectomy. We applied hierarchical regression analysis retrospectively to examine the effect of 3 predictor variables on middle cerebral artery velocity (MCAV): age at follow-up TCD (step 1); MCAV at baseline (preoperative in surgical cases: step 2); and surgical status (removal of tonsils or not; 0, 1: step 3). Age at follow-up TCD significantly predicted follow-up MCAV ($F_{1,153} = 28.14$ [P < .001]; $R^2 = .16$). When baseline MCAV was added (step 2), the explanatory power of the model significantly increased ($F_{2,152} = 57.87$ [P < .001]; $\Delta R^2 = .28$ [P < .001]). However, at step 3, surgical status did not add to the power of the model ($F_{1,151} = 38.45$ [P < .001]; $\Delta R^2 = .39$ [P = .656]). As can be seen from the summary in Table 1, older age was predictive of lower MCAV at follow-up, whereas higher baseline MCAV was predictive of higher follow-up MCAV. Surgical status did not affect final MCAV levels. Of 18 children with MCAV of ≥200 cm/second, 15 had a repeat TCD measurement, of whom 2 had had adenotonsillectomy between the measurements; MCAV decreased from 262 to 250 cm/second in one patient and from 212 to 205 cm/second in the other; both still fell into a range that, according to Stroke Prevention in Sickle Cell Anemia (STOP) criteria, would make transfusion required. Of 17 children with MCAV of ≥170 but <200 cm/second (conditional), all had repeat TCD measurements, of whom 2 had had adenotonsillectomy between the measurements; MCAV decreased from 179 to 170 cm/second in one and from 194 to 14 cm/second in the other.

Although these data suggest that adenotonsillectomy has limited influence on MCAV in children with SCA, surgery may be performed for different reasons, including obstructive sleep-disordered breathing, chronic or intermittent oxyhemoglobin desaturation, or, alternatively, chronic infection. In addition, adenotonsillectomy for patients with SCA is often undertaken some time after the onset of symptoms because of operative risk factors. Postoperative MCAV, therefore, may be differentially affected by the original etiology and the duration and severity of exposure to infection, inflammation, and hypoxemia as well as the treatment per se. Perioperative carotid dissection is also a risk; this was a possible cause of the very low postoperative velocity in 1 of our conditional patients. Once initiated, cerebrovascular disease in SCA is rarely completely reversible, but it is possible that in the cases reported by Bader-Meunier et al, a detrimental effect of hypoxemia and chronic infection on the endothelium was reversed before secondary vascular changes commenced.

Polysomnography would be informative for future study of young children with SCA to clarify the associations between adenotonsillar hypertrophy, objectively rated obstructive sleep-disordered breathing, and increased MCAV. This would be comparable with the design of our original study in children without SCA.

### Table 1: Summary of Hierarchical Regression Analysis for Variables Predicting MCAV (N = 155)

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE</th>
<th>β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at follow-up MCAV</td>
<td>−4.04</td>
<td>.76</td>
<td>−.39</td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at follow-up MCAV</td>
<td>−3.03*</td>
<td>.64</td>
<td>−.30</td>
</tr>
<tr>
<td>Baseline MCAV</td>
<td>0.55*</td>
<td>.06</td>
<td>.54</td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at follow-up MCAV</td>
<td>−2.97*</td>
<td>.66</td>
<td>−.29</td>
</tr>
<tr>
<td>Baseline MCAV</td>
<td>0.55*</td>
<td>.06</td>
<td>.53</td>
</tr>
<tr>
<td>Surgical status (0 = no surgery)</td>
<td>−2.88</td>
<td>.64</td>
<td>−.03</td>
</tr>
</tbody>
</table>

*P < .01.

---

**In Reply.**

We appreciate the possible implications for children with sickle cell anemia (SCA) in the response by Bader-Meunier et al to our article on increased cerebral blood flow velocity in children with SCA with mild sleep-disordered breathing. Although there seems to be a link between sleep-disordered breathing and changes to the cerebral vasculature, the mechanism has not been elucidated. Candidates include the mechanical effects of snoring, associated hypoxemia, and infection/inflammation. Tonsillectomy is sometimes performed for recurrent tonsillitis or relief of upper airway obstruction in children with SCA, although mean overnight oxyhemoglobin saturation may not improve, and there are few systematically collected data that have addressed a relationship with SCA complications such as transcranial Doppler (TCD) abnormality.

TCD testing has been undertaken at least once in 384 children with homozygous SCA in the London cohort, and adenotonsillectomy has been performed on clinical grounds. In 183 (47.7%) patients, TCD data are available on 2 occasions (mean age at follow-up: 11.2 years [SD: 4.5 years]). Of these, 40 (21.9%) had undergone adenotonsillectomy between the 2 TCD measurements (mean age at surgery: 7.7 years [SD: 3.9 years]), 28 had adenotonsillectomy before the first TCD (not included in this analysis), and 115 had not undergone adenotonsillectomy. We applied hierarchical regression analysis retrospectively to examine the effect of 3 predictor variables on middle cerebral artery velocity (MCAV): age at follow-up TCD (step 1); MCAV at baseline (preoperative in surgical cases: step 2); and surgical status (removal of tonsils or not; 0, 1: step 3). Age at follow-up TCD significantly predicted follow-up MCAV ($F_{1,153} = 28.14$ [P < .001]; $R^2 = .16$). When baseline MCAV was added (step 2), the explanatory power of the model significantly increased ($F_{2,152} = 57.87$ [P < .001]; $\Delta R^2 = .28$ [P < .001]). However, at step 3, surgical status did not add to the power of the model ($F_{1,151} = 38.45$ [P < .001]; $\Delta R^2 = .39$ [P = .656]). As can be seen from the summary in Table 1, older age was predictive of lower MCAV at follow-up, whereas higher baseline MCAV was predictive of higher follow-up MCAV. Surgical status did not affect final MCAV levels. Of 18 children with MCAV of ≥200 cm/second, 15 had a repeat TCD measurement, of whom 2 had had adenotonsillectomy between the measurements; MCAV decreased from 262 to 250 cm/second in one patient and from 212 to 205 cm/second in the other; both still fell into a range that, according to Stroke Prevention in Sickle Cell Anemia (STOP) criteria, would make transfusion required. Of 17 children with MCAV of ≥170 but <200 cm/second (conditional), all had repeat TCD measurements, of whom 2 had had adenotonsillectomy between the measurements; MCAV decreased from 179 to 170 cm/second in one and from 194 to 14 cm/second in the other.

Although these data suggest that adenotonsillectomy has limited influence on MCAV in children with SCA, surgery may be performed for different reasons, including obstructive sleep-disordered breathing, chronic or intermittent oxyhemoglobin desaturation, or, alternatively, chronic infection. In addition, adenotonsillectomy for patients with SCA is often undertaken some time after the onset of symptoms because of operative risk factors. Postoperative MCAV, therefore, may be differentially affected by the original etiology and the duration and severity of exposure to infection, inflammation, and hypoxemia as well as the treatment per se. Perioperative carotid dissection is also a risk; this was a possible cause of the very low postoperative velocity in 1 of our conditional patients. Once initiated, cerebrovascular disease in SCA is rarely completely reversible, but it is possible that in the cases reported by Bader-Meunier et al, a detrimental effect of hypoxemia and chronic infection on the endothelium was reversed before secondary vascular changes commenced.

Polysomnography would be informative for future study of young children with SCA to clarify the associations between adenotonsillar hypertrophy, objectively rated obstructive sleep-disordered breathing, and increased MCAV. This would be comparable with the design of our original study in children without SCA.

---

**Alexandra M. Hogan, PhD**
Developmental Brain-Behaviour Unit

**Catherine M. Hill, BM, FRCPCH**
Clinical Neurosciences

**Romola Bucks, PhD**
School of Psychology
University of Southampton
Southampton SO17 1BJ, United Kingdom

**Paul Telfer, DM, MRCP, FRCPath**
Department of Haematology

---

**LETTERS TO THE EDITOR**
REFERENCES


Do We Need to Incorporate Pharmacogenetics in Randomized, Controlled Trials of Frequently Used Medicines?

*To the Editor.—*

Clark et al1 reported that in children 6 to 17 years old with pain from a musculoskeletal injury, ibuprofen (10 mg/kg) provided the best analgesia as compared with acetaminophen (15 mg/kg) and codeine (1 mg/kg). In their well-designed randomized, controlled trial, they used the standard doses of these 3 medications. However, it is important to note that although ibuprofen was more efficacious in providing adequate analgesia, only 52% of the children in this group could be defined as receiving “adequate analgesia,” which left 48% of the children with “inadequate analgesia.” Is this because the wrong dose was used, or does this reflect interindividual variation in pharmacokinetics and/or pharmacodynamics of the investigated medications? As a consequence, the question arises of whether the incorporation of pharmacogenetic analysis of drug-metabolizing enzymes, transporters, and/or receptors involved in the pharmacokinetics/pharmacodynamics of these drugs might result in more individualized adequate analgesia. In the case of ibuprofen, there is compelling information2 that mutations in cytochrome 2C8 (CYP2C8) significantly change the clearing capacity of the individuals who are heterozygous or homozygous mutants for this drug-metabolizing enzyme. Perhaps these individuals are represented by the children in this study who received adequate analgesia because they were just exposed longer to higher concentrations. I would like to hear from the authors how many children reached adequate analgesia in the codeine and acetaminophen groups. Especially the use of codeine, which is a prodrug and needs metabolism to morphine using cytochrome 2D6 (CYP2D6), clearly depends on the CYP2D6 phenotype of the children exposed to this drug.3

In general, I believe that in the (near) future incorporation of a pharmacogenetic component in randomized, controlled trials might improve the clinical usefulness of the results of these state-of-the-art investigations. Ideally, one would like to randomly assign children who have the same genetic background as it relates to drug-metabolizing enzymes, transporters, and receptors to be able to really decipher the differences between the drugs used for pain relief.

*John N. van den Anker, MD, PhD*

Department of Pediatrics

George Washington University

Washington, DC 20010

REFERENCES


In Reply.—

We agree with Dr van den Anker’s concern that although ibuprofen was more efficacious in providing adequate analgesia, only 52% of the children in this group could be defined as receiving adequate analgesia (visual analog scale <30 mm) at 60 minutes. As shown in our Table 2, 40% (95% confidence interval: 31%–50%) of children in the codeine group and 36% (95% confidence interval: 27%–46%) of children in the acetaminophen group were defined as receiving adequate analgesia at 60 minutes. The high level of inadequate pain control among all 3 study groups brings into question other strategies to relieve pain, which may include larger doses of each medication, the addition of a second medication,
members may suffer a significant burden of anxiety that impacts quality of life. Although venom immunotherapy improves quality of life, the recommendation that anaphylaxis preparedness include an indefinite prescription for self-injectable epinephrine in venom-allergic children with strictly cutaneous symptoms may impair quality of life. For patients with a low risk of progressive anaphylaxis (such as children with strictly cutaneous reactions to stinging insects), a process of shared decision-making may be appropriate, because significant trade-offs are likely to exist between preparedness, inconvenience, and anxiety. It seems that, on the basis of the available evidence, the decision to prescribe prophylactic self-injectable epinephrine for children with generalized acute urticaria after an insect sting should be an individualized decision, not a universal recommendation.

In Reply.—

We thank Drs Gaines and Shaker et al for their thoughtful comments. We fully agree that it is not always easy to assess the clinical risk of future anaphylaxis, which influences the decision to prescribe self-injectable epinephrine.

We overstated the risk of subsequent anaphylaxis in children with generalized acute urticaria (mild cutaneous systemic reactions) after an insect sting as ~10%. It is, in fact, precisely 6.74% as stated in the study by Golden et al, which we cited at the end of our clinical report. It is important to note that this study by Golden et al included the earlier data published by Valentine et al, which are described in the letters by Drs Gaines and Shaker et al.

In the study by Golden et al, follow-up data were obtained on 89 children with initial mild cutaneous systemic reactions who had not received venom immunotherapy. When subsequently stung, 87% of the children had no systemic reaction, 6.74% had another mild cutaneous systemic reaction, 6.74% had a moderate systemic allergic reaction, and none had a severe allergic reaction. A moderate systemic reaction was defined as “signs and symptoms of a cutaneous reaction as well as discomfort in the throat or chest, mild symptoms of airway obstruction, light-headedness, and dizziness or mild hypotension,” a description that is consonant with the current definition of anaphylaxis.

In this context, it is important to note that physicians cannot assume that children and their caregivers necessarily recognize and report all anaphylaxis symptoms, because even trained health care professionals underrecognize this disorder.

Therefore, we continue to recommend that pediatricians prescribe self-injectable epinephrine for children who have mild cutaneous systemic reactions consisting of generalized acute urticaria after an insect sting, as recommended in a recent practice parameter. Some, like Shaker et al, may consider the prescription optional given the risks. At the very least, prescription of self-injectable epinephrine should be discussed with the caregivers who are responsible for the child, and consultation with an allergist should be considered.

REFERENCES


doi:10.1542/peds.2007-0731
Ventilatory Pump Failure and Strategies to Prevent Bronchopulmonary Dysplasia

To the Editor.—

It was with great interest that I read Dr Aly’s commentary regarding strategies for preventing bronchopulmonary dysplasia (BPD).1 The avoidance of intubation and conventional mechanical ventilation is a commendable goal when clinically feasible, and more widespread effective use of early nasal continuous positive airway pressure (CPAP) holds potential clinical promise to reduce the incidence and severity of BPD. However, the effectiveness of any CPAP strategy will depend on factors that extend beyond operator experience and include, among other physiologic variables, the severity of the infant’s underlying parenchymal lung disease and the maturity of the newborn’s ventilatory pump. Although the former is widely appreciated as an important respiratory variable in predicting the need for conventional mechanical ventilator support, the latter is often overlooked, to the detriment of establishing clinically effective respiratory support for the extremely low birth weight (ELBW) neonate. This neonatologist’s experience suggests that ventilatory pump failure is often an indication for initiation of conventional mechanical ventilator support in early postnatal life and a contributor to the duration of such support as well, especially in the ELBW preterm neonate.2 Immaturity of the ventilatory pump in preterm neonates is characterized by diminished central neural respiratory drive, contractile properties of the diaphragm and accessory muscles of respiration that limit force generation, and thoracoabdominal mechanics that place the preterm neonate at a disadvantage when challenged with lung disease.3,4 These factors individually and collectively predispose to ventilatory pump failure in ELBW neonates.3,4 Despite knowledge of these deficits, little is known regarding how to enhance the effectiveness of the preterm newborn’s ventilatory pump. The use of caffeine to increase central neural respiratory drive is currently the only widely used intervention aimed at enhancing the effectiveness of the preterm neonate’s ventilatory pump, and it is of interest that caffeine therapy for apnea of prematurity was shown recently to reduce the rate of BPD in very low birth weight newborns.5 Novel noninvasive efforts designed to improve chest wall stability and support lung volume in preterm newborns may prove beneficial but await more widespread application in the clinical arena.6 Until additional strategies are developed to enhance central neural respiratory drive and newborn respiratory muscle contractile properties and improve the preterm neonate’s chest wall mechanics, success in using CPAP in the respiratory support of the ELBW neonate will be suboptimal and its therapeutic potential to reduce the incidence of BPD not fully realized.

Jon F. Watchko, MD
Division of Newborn Medicine
Department of Pediatrics
University of Pittsburgh School of Medicine
Pittsburgh, PA 15213

REFERENCES

In Reply.—

I am grateful for the letter from Dr Watchko, which stresses the importance of neuronal respiratory drive, contractile properties of respiratory muscles, and thoracoabdominal mechanics, in addition to parenchymal lung disease in premature infants to improve the experience with the use of continuous positive airway pressure (CPAP). Of course, optimizing all these factors will help infants to breathe easier and avoid the need for tracheal intubation and mechanical ventilation. However, it is important to mention that the benchmark incidence of bronchopulmonary dysplasia (BPD) in very low birth weight infants is 29%, whereas in our recent
experience with CPAP the incidence of BPD was only 6%. Similar figures for BPD have been published for the CPAP experience at Columbia University. Therefore, the use of the currently available CPAP is vastly successful. Our goal is to focus on making these BPD figures universally replicable in all units. Once this milestone is established, efforts should then be focused on optimizing the “ventilatory pump” to further lower BPD to <6%. At the current stage, I find it distracting to focus on different medications and ventilatory approaches while the available CPAP technique can be feasible and successful in our population. Subsequent efforts that can support premature lungs and/or ventilatory pump should be addressed only after passing this first important milestone.

Hany Aly, MD
Department of Newborn Services
George Washington University Hospital and
Children’s National Medical Center
Washington, DC 20037

REFERENCES

doi:10.1542/peds.2007-1163

Newborn Screening for Cystic Fibrosis in New York State

To the Editor.—

We wish to rectify numerous inaccuracies in the recent article by Giusti et al.1 Although there is no “state coordinator” per se for cystic fibrosis (CF), a follow-up CF team does exist to provide assistance to anyone who contacts the screening program. The authors asserted that an immunoreactive trypsinogen value was unavailable for 1 cited infant; however, to our knowledge, no inquiry was made to the program in that individual case. Section 69–1.7 of the New York State Public Health Law 2500a specifies the roles of specialized care centers, also known as treatment centers. Status as a designated treatment center is voluntary, not mandated. With respect to reimbursement for genetic counseling services, New York State has the Genetic Services Program, which awards monies so that every state resident has access to genetic counseling services. Only 1 CF center lacks an award through this program.

The statement that the program will not accept a second specimen was erroneous. In fact, we receive >15,000 “repeat” specimens annually, corresponding to ~6% of New York State births. The program requests some of these specimens if either the original sample provided a borderline value or it was unsuitable for analysis because of poor/early collection. The rest, flagged as unsolicited repeats, are screened for the entire panel unless otherwise indicated. All repeat specimens are linked to the original samples in our computer system, and we request the results of any subsequent testing performed so that we can close those cases and ensure that all infants received the medical attention to which they are entitled under the law.

Last, the statement that “the 5/7/9T variant is not a reflex test in the New York state protocol” was incorrect. Screening began on October 7, 2002, and every report of an R117H variant includes the intron-8 polymorphism status.2,3 At the time of the original publication,1 reflex testing for 3199del6 was under regulatory review; it has since been implemented in New York State.

The general strategy of the program has been to proceed conservatively until sufficient data have been collected to warrant changes. For example, we began medium-chain acyl-coenzyme A dehydrogenase deficiency screening in 2002. After review of our data, we reduced the number of referrals from 73 in 2003 to 41 in 2005, a 40% reduction. Similar changes to the CF screening protocol are forthcoming in New York State.

Michele Caggana, ScD, FACMG
Laura Helton, MD, MPH
New York State Newborn Screening Program
Wadsworth Center
New York State Department of Health
Albany, NY 12201

REFERENCES

doi:10.1542/peds.2007-1048

In Reply.—

We are happy to reply to the letter to the editor submitted by Drs Caggana and Helton. The purpose of writing our article was to report on cystic fibrosis (CF) newborn screening in a state with an ethnically diverse population, to make recommendations to improve the sensitivity of the screening algorithm, and to help physicians in states that are in the planning stages of a CF newborn
screening program to address some of the problems that we encountered.

In reference to the absence of a “state coordinator,” we fully agree that the Wadsworth Center has a well established team to provide assistance related to the newborn screening program. However, in our article, we were referring to the fact that there is no funding to support a coordinator to facilitate the referral of screen-positive infants to regional CF care centers, where a diagnostic sweat test can be performed. Other states, such as Massachusetts, have developed funding mechanisms to support such a position, which has facilitated making contact with screen-positive infants and assuring timely sweat testing.\(^1\) It is our opinion that developing a funding source for such a position would be an effective intervention to help resolve the large number of cases that have been lost to follow-up in New York. We recommend that any state that is in the planning stage of a CF newborn screening program consider the benefit of establishing such a position.

With regard to reimbursement for genetic counseling services, we acknowledge that the New York State Department of Health provides funding to many hospitals in the state for genetic counseling services. However, the CF care centers, where the bulk of the counseling and coordination for follow-up of positive CF-screened infants occurs, do not receive any funding to help defer the cost of additional time committed by staff members to ensure that the screening program runs effectively. Other states, such as New Jersey (R. Zanni, MD, personal communication, April 2007), have been successful in acquiring additional funding for the CF care centers to help pay for the manpower needs of their screening program. Once again, we recommend that any state that is in the planning stage of a CF newborn screening program consider a mechanism to generate additional funds for this purpose.

We acknowledge that the Wadsworth Center performs repeat specimen testing. However, in our article we raised the issue of developing a mechanism to accept a repeat immunoreactive trypsinogen (IRT) result at 2 weeks of age to close a case without the need for a referral for a sweat test. The screening program in Colorado collects a repeat specimen obtained at 2 weeks of age. We believe that accepting a repeat IRT at 2 weeks of age could be a solution to the high false-positive rate that we have experienced in New York State during the first 2½ years of screening for CF.\(^2\) The large number of screen-positive infants with an elevated IRT and no CF mutations detected would be suitable for requesting repeat IRT testing.

We acknowledge that currently in New York State the intron-8 polymorphism is reported whenever the R117H mutation is detected. However, the issue that we raised in our article was whether 7T/9T polymorphisms should be reported as screen-positive. In view of the high frequency of R117H-7T identified by CF newborn screening, the uncertain outcome of the asymptomatic children, and physicians’ difficulty in managing these situations, Scotet et al\(^3\) proposed the withdrawal of the R117H variant from the panels of CF transmembrane conductance regulator (CFTR) mutations used in CF newborn screening. Any state that is planning a CF newborn screening program must be aware that “[u]se of mutations associated with milder phenotypes may lead to results that are difficult to interpret.”\(^4\) We acknowledge that this is a controversial issue that needs to be considered by states that are in the planning stages of a CF newborn screening program.

We fully agree with the general strategy of proceeding in a conservative manner; however, we support the recommendation of the Centers for Disease Control and Prevention report on newborn screening for CF: “states that implement newborn screening for CF should collect follow-up data in collaboration with CF care centers and analyze this information to monitor and improve the quality of CF newborn screening.”\(^5\) Such an annual-review mechanism would facilitate making adjustments to the screening algorithm with the goal of improving the sensitivity of the program. We are happy to report that in New York State we currently have monthly conference calls and have annual meetings planned.

---

REFERENCES
5. Grosse SD, Boyle CA, Botkin JR, et al. Newborn screening for cystic fibrosis: evaluation of benefits and risks and recommen-
Prenatal Chlorpyrifos and Early Neurodevelopment: How Good Is the Science?

To the Editor.—

In their article, Rauh et al1 drew many false conclusions and implicated prenatal exposure to chlorpyrifos as the cause of deficits in Mental Developmental Index and Psychomotor Development Index 3 years later. The conclusions were false because:

1. The mean Mental Developmental Index scores (see their Table 1) between the high- and low-exposure groups are clinically meaningless: 94 and 94 at 12 months; 84 and 85 at 24 months; and 87 and 90 at 36 months.
2. The average high- and low-exposure Psychomotor Development Index scores were 93 and 97 at 12 months, 99 and 97 at 24 months, and 96 and 102 at 36 months; and measured IQ levels ranged between 9 and 30 points (profound mental retardation [MR]: 0–29; severe MR: 20–29; moderate MR: 30–49; mild MR: 50–69; dull normal: 70–89; normal: 90–110; bright normal: 111–119; superior: 120–139; very superior: 140 to ≥160).
3. With no standards defining “high” and “low” exposure, the authors concluded falsely that 20.6% of the sample had high exposure to chlorpyrifos.
4. The authors tried to eliminate maternal education bias by dichotomizing maternal education into high school graduate or nongraduate. This dubious maneuver masked, but did not eliminate, educational bias. The authors should have used a better index of socioeconomic status, such as family income.
5. The sample mean was substituted for intelligence scores that were missing for 29 women. Of 14 methods of data imputation, Engels and Diehr2 concluded that: “Imputations that used no information specific to the person, such as using the sample mean, had the worst performance.”
6. A correlation of 0.76 means that the variability in umbilical cord samples explains only 0.76^2 or 58% of the variability in maternal plasma, which indicates that the authors’ substitution of 1 variable for the other is not valid.
7. The high- and low-exposure groups (see their Table 1) were grossly different in race/ethnicity characteristics: the high-exposure group contained 24% black and 15% Dominican mothers, whereas the lower-exposure group was comprised of 76% black and 85% Dominican mothers. These huge differences inextricably confound race/ethnicity with exposure level and, therefore, served to negate the role of chlorpyrifos as an etiologic factor for later deficits in psychomotor and cognitive performance.
8. The finding that high-exposure children had more attention-deficit/hyperactivity disorder problems is also meaningless, because, as the authors themselves concluded, the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition “has low sensitivity for assessing the inattentiveness of preschool-aged children.”

Domenic V. Cicchetti, PhD
Child Study Center
Departments of Psychiatry and Biometry
Yale School of Medicine
New Haven, CT 06512

REFERENCES

In Reply.—

We reported that inner-city children who were highly exposed in utero to the insecticide chlorpyrifos were significantly more likely than lower-exposed children to experience motor and mental delays and behavior problems related to attention, attention-deficit/hyperactivity, and pervasive developmental disorders.

In response to Dr Cicchetti’s remarks:

1. We strongly disagree with Dr Cicchetti’s claim that the significant chlorpyrifos effect on mental development was “clinically meaningless.” A Bayley developmental score of ≥85 prompts referral to early intervention services and is far from trivial.1 Exposures that produce small shifts in the mean often result in more children who meet diagnostic criteria, which is an outcome that is of great concern to clinicians.

In reply —

We reported that inner-city children who were highly exposed in utero to the insecticide chlorpyrifos were significantly more likely than lower-exposed children to experience motor and mental delays and behavior problems related to attention, attention-deficit/hyperactivity, and pervasive developmental disorders.

In response to Dr Cicchetti’s remarks:

1. We strongly disagree with Dr Cicchetti’s claim that the significant chlorpyrifos effect on mental development was “clinically meaningless.” A Bayley developmental score of <85 prompts referral to early intervention services and is far from trivial.1 Exposures that produce small shifts in the mean often result in more children who meet diagnostic criteria, which is an outcome that is of great concern to clinicians.

2. Contrary to Dr Cicchetti’s point, we used a clearly stated, a priori standard for high chlorpyrifos exposure, defined as ≥6.17 pg/g, based on a previous report of reduced birth weight among children with exposures above this level.2

3. We do not understand why Dr Cicchetti thinks that using high-school degree to adjust for maternal educational level is a “dubious maneuver” and note his own use of this variable in an analysis of cocaine effects on development.3 Because our sample was uniformly low income, education was the preferred covariate for social class.

PEDIATRICS Volume 120, Number 1, July 2007 243
4. Maternal intelligence, although controlled, was not significant in our analysis.

5. Because concentrations of chlorpyrifos in maternal blood and/or adipose tissue are in steady state,\textsuperscript{4} cord-blood level provides a very reasonable dosimeter for the amount of parent compound transferred from mother to fetus. The actual level was derived from the regression of cord-blood levels on maternal blood levels, as clearly described, and was not a “substitution.”

6. We not only included race/ethnicity in all models to control for confounding, but we also used stratified analyses, showing a significant chlorpyrifos effect within each ethnic group, independent of race, and not “inextricably confounded with race/ethnicity.”

7. Because attention-deficit/hyperactivity disorder is hard to diagnose in preschool-aged children, we used Achenbach’s Child Behavior Checklist to assess behavior problems rather than make diagnoses.

To conclude, in 2001 the Environmental Protection Agency banned residential, not agricultural, use of chlorpyrifos because of its potential neurotoxic consequences for children. Our article provided evidence that supports that regulatory decision.

Virginia A. Rauh, ScD
Heilbrunn Center for Population and Family Health
Mailman School of Public Health
Columbia University
New York, NY 10032

Robin W. Whyatt, DrPH
Frederica P. Perera, DrPH
Howard F. Andrews, PhD
Robin Garfinkel, PhD
Deliang Tang, DrPH
Center for Children’s Environmental Health
Mailman School of Public Health
Columbia University
New York, NY 10032

REFERENCES


doi:10.1542/peds.2007-1198

Pediatric Palliative Care

To the Editor—

As a nurse working in an intensive care nursery, I have dealt with children who have chronic illnesses and sometimes face death. This is why I found the article by Feudtner et al entitled “Hopeful Thinking and Level of Comfort Regarding Providing Pediatric Palliative Care: A Survey of Hospital Nurses” very enlightening.\textsuperscript{1} I have learned about palliative care through experience and from our hospital’s Pediatric Advanced Comfort Care Team (PACCT). There are still situations in which I am not comfortable speaking about the death of a child, which is why the level of comfort that nurses have regarding palliative care is extremely important. I commend your efforts in providing awareness about pediatric palliative care and agree with the authors’ conclusion that many nurses feel proficient regarding pain management and inadequate when talking with children and families about dying. This article showed that there is a need for more education in patient care settings regarding palliative care. By increasing knowledge related to palliative care, a positive effect could be obtained for both patients/families and nurses.

Feudtner et al found that “[n]urses’ level of hope is associated with their self-reported comfort and competence regarding palliative care.” To increase comfort levels of nurses, hospitals should be encouraged to implement courses like the End-of-Life Nursing Education Consortium, which was developed in September 2001. The End-of-Life Nursing Education Consortium is a training course that informs nurses about communication skills, ethics, cultural considerations, pain/symptom management, care at the time of death, grief, and bereavement.\textsuperscript{2} Hospitals should also be encouraged to create a pediatric supportive care team like the one my hospital, Children’s Mercy, implemented (the Pediatric Advanced Comfort Care Team). These teams can enhance the “quality of life in the face of an ultimately terminal condition.”\textsuperscript{3}

New research and education can provide nurses with useful information about palliative care. Interventions should be studied in clinical trials to determine the effectiveness in achieving better patient, family, and health care provider outcomes.\textsuperscript{1} Because nurses have a distinctive role in promoting and providing care for terminal children and their families, education and teamwork are imperative to provide pediatric nurses the confidence that is needed in these situations. I hope that \textit{Pediatrics} will continue to support the efforts of palliative care education and look forward to reading more literature regarding this important topic.
Choosing the Best Practice: Evidence to Support Fluoroquinolone Drops for Acute Otitis Media Through Tympanostomy Tubes

To the Editor.—

In his commentary, “Why Don’t Those Ear Drops Work for My Patients?” Isaacson presented his argument against the use of ototopical fluoroquinolones alone or in combination with a steroid as the first line of defense for the treatment of acute tube otorrhea in infants. He made several points to discount the use of these agents in pediatric patients with acute otitis media with tympanostomy tubes. However, rooted in the first sentence of the commentary was the fact that seems to lie at the epicenter of his concerns: “Topical Ciprofloxacin/Dexamethasone Is Superior to Oral Amoxicillin/Clavulanic Acid in Acute Otitis Media With Otorrhea Through Tympanostomy Tubes,” [is] the latest in a series of pharmaceutical-industry–funded articles. 

Isaacson’s first critique of this article pertains to the study population. He stated that “[i]f one were trying to design a study to demonstrate the superiority of ototopical drops over oral antibiotics, the ideal population would be older children.” This statement directly contradicted his other criticism that “more than half of their subjects were infants with ear drainage for more than 1 week or who had been treated with antibiotics should have also been the focus of the study.” Furthermore, experts agree that otitis media studies with a preponderance of infants are more compelling, because these are the patients who are most difficult to treat.

Isaacson continued by stating that patients with ear drainage for more than 1 week or who had been treated with antibiotics should have also been the focus of the study.

The goal of this study was to look at a cohort of patients with untreated acute otitis media and compare topical antibiotic therapy to systemic therapy with the antibiotic that most guidelines consider the “best” for treating otitis media. Evaluating a different strategy of treatment supports the new directive of the Centers for...
Disease Control and Prevention to reduce antimicrobial resistance through the promotion of more appropriate antibiotic use. Contrary to this goal, this commentary both directly and indirectly supported systemic paradigms, although topical antibiotics have been shown to be safe and effective and less likely to lead to increased antibiotic resistance. Isaacson referred to a “protocol for controlling acute tube otorrhea . . . devised by clinicians at the Otitis Media Research Center in Pittsburgh, Pennsylvania.” However, the reference cited to support the protocol is not one directly published by the Otitis Media Research Center in Pittsburgh but was written by Isaacson and Richard Rosenfeld. It was extracted from a book with a copyright suggesting that the article was written before the availability of most of the supporting evidence for ototopical quinolone drops.

Isaacson credited our exclusion of children from whom pure cultures of *Pseudomonas aeruginosa* were isolated. However, we disagree with his criticism that the “organisms they recovered from culture do not reflect the usual pathogens of acute otitis media, with only 15 of 154 cultures growing *Streptococcus pneumoniae*.” Having participated in numerous studies with a collective experience of >1000 patients, we have found that this same microbiological profile has almost always been the case. Furthermore, the data on the recovery rates of *S pneumoniae* must be evaluated against the rates after pneumococcal vaccination. Before this time, Roland et al. reported a rate of *S pneumoniae* recovery of 17% during 2000–2001. On the basis of an approximate 30% reduction in any *S pneumoniae* isolate in acute otitis media as reported by the Finnish Otitis Media Study Group, the 10% recovery rate for *S pneumoniae* in our study was expected.

We also disagree that drops instilled into the ear canal do not enter the middle ear. In addition to pharmacokinetic studies, which challenge this statement, one would not expect cure rates over double expected with spontaneous resolution (41%) after 1 week of treatment. In addition, no studies to our knowledge have ever suggested that drops should be “dribbled on top of a collection of pus” as suggested in this commentary. Aural toilet is recommended by most published authorities on this subject, including the World Health Organization. Dry mopping has been shown to be an effective means of aural toilet, and this clearly is practical in any physician’s office.

Finally, Isaacson made a statement that we enthusiastically embrace: “We should think twice about our approach to infectious disease in children and choose best practices on the basis of strong evidence.” We believe that strong evidence exists, and more is being generated to support both the safety and efficacy of ototopical quinolones and quinolone/steroid combinations such as ciprofloxacin/dexamethasone as first-line treatment for tympanostomy tube otorrhea. Without question, several of the studies have been funded by corporate sponsors with no attempt to hide or conceal this fact. We should appreciate such funding support, because it is extremely unlikely that noncorporate sponsors such as the National Institutes of Health would fund such work, especially in light of federal budget cuts. We do, however; recognize the importance of constructive and ethical partnerships between medicine and industry to advance the care of our patients. Although improprieties regarding the conduct of certain industry-funded studies have been a reality, it is unwarranted and erroneous to presume guilt by association for all such studies. When one reviews the type of evidence that Isaacson expects, such as Cochrane reviews, meta-analyses, and evidence-based, best-practice guidelines, one finds no such exclusion of evidence on the basis of the funding source. In many ways, industry-funded studies undergo more careful oversight and scrutiny. They must withstand preapproval by the US Food and Drug Administration and must be approved by several investigators and subinvestigators, as well as countless institutional review boards. They are subject to internal reviews, external audits, and, in select cases, scientific advisory boards convened by the Food and Drug Administration. And, like all studies, they can only be published after thorough peer review. We should be no more willing to condemn all such industry-sponsored trials than we are willing to condemn all physicians for the documented malfeasances of a few.
not be the right drug for every clinical situation, specifically for severe otorrhea in a febrile infant.4

As physicians entrusted with the welfare of young children, we have a responsibility to find the safest, most effective treatments. Without specific strong evidence, we should hesitate to abandon proven cures for potentially serious illnesses.

Glenn Isaacson, MD, FAAP
Department of Otolaryngology, Head and Neck Surgery
Temple University School of Medicine
Philadelphia, PA 19140

REFERENCES


doi:10.1542/peds.2007-1268

Neonatal Blue-Light Phototherapy Could Increase the Risk of Dysplastic Nevus Development

To the Editor.—

In their letter, Csoma et al1 described the possible concern that blue-light phototherapy would increase nevus development. They screened 747 school-aged children in their teens and reported that 44.6% of these children had received phototherapy. Although there were no differences in the prevalence of melanocytic nevi, the children exposed to blue-light phototherapy had higher numbers of moles and had a significantly higher prevalence of atypical nevi. Because atypical nevi are an important risk factor for the development of malignant melanoma, the authors raised the concern that phototherapy with blue lamps could have significant implications in the development of atypical nevi.

Although this was an interesting observation that disputes a previous report,2 the data are somewhat difficult to interpret, and the conclusions and statistical validity are not completely evident. We were not given sufficient information about patient demographics or how the data collection was undertaken. For a parent to remember whether their child was treated with blue-light phototherapy 14 to 18 years later would be very

REFERENCES


doi:10.1542/peds.2007-1268

In Reply.—

Ciprofloxacin is an excellent antibiotic, and dexamethasone is a glucocorticoid with proven effect. The problem arises when 2 middle-aged drugs are combined into a new preparation that, at $14 per gram,1 is more valuable than gold or heroin.

Like Dohar et al and the employees of Alcon Laboratories, Inc, I long for a time when research funds are abundant and fine physician-scientists are free to design clinical experiments with nothing but the quest for truth in mind. Shareholder pressure to realize profit after the large investment of new product development has forced manufacturers to present perfectly fine medications as panaceas, often with tragic results.2,3

What inspired the forcefully worded letter from Dohar et al? It was the suggestion that ciprodex might

Krista Crenshaw, MS
Alcon Research, Ltd
Fort Worth, TX 76130

Ramzi Younis, MD
Department of Pediatric Otolaryngology
Leonard M. Miller School of Medicine
Bascom Palmer Eye Institute
Miami, FL 33136

Michael Poole, MD
Savannah, GA 31406

REFERENCES


doi:10.1542/peds.2007-1268

In Reply.—

Ciprofloxacin is an excellent antibiotic, and dexamethasone is a glucocorticoid with proven effect. The problem arises when 2 middle-aged drugs are combined into a new preparation that, at $14 per gram,1 is more valuable than gold or heroin.

Like Dohar et al and the employees of Alcon Laboratories, Inc, I long for a time when research funds are abundant and fine physician-scientists are free to design clinical experiments with nothing but the quest for truth in mind. Shareholder pressure to realize profit after the large investment of new product development has forced manufacturers to present perfectly fine medications as panaceas, often with tragic results.2,3

What inspired the forcefully worded letter from Dohar et al? It was the suggestion that ciprodex might
difficult. If information was based on old records, it is also not clear whether all charts specifically indicated whether a patient was under blue-light versus white-fluorescent-light therapy. Additional important confounders besides phototherapy, such as family history and sun exposure, were not discussed. In addition, 44.6% of the population received phototherapy. This is a very high number compared with what one would expect in a typical newborn population. Therefore, the question of which infants were included in the cohort becomes more important. It is also not clear how the reported odds ratios were derived. Was this a multivariate analysis, or were the odds ratios just obtained from a 2-by-2 table? The lack of information provided in this letter makes it difficult to raise concern about the risks of phototherapy in the pediatric community. Solid evidence with carefully conducted statistics and careful explanations about the cohort are needed. Phototherapy is an important tool in the management of neonatal jaundice. When severe, this condition may result in kernicterus. Although this event is rare in occurrence, it has devastating consequences that must not be forgotten. Phototherapy has been used since 1958. There should be large cohorts to assess whether the concerns for the development of atypical nevi are real or only applicable to the subset of children described in this letter.

We need to remember the devastating consequences of our reduced vigilance for hyperbilirubinemia in the late 1980s and early 1990s. We must seriously weigh the resurgence of kernicterus against the potential for moles and nevi until more strategies are available to prevent hyperbilirubinemia.

Rather than a letter, the investigators should have provided a full manuscript with sufficient information to address all of the concerns raised regarding the methodology, the cohort, and family history.

Phyllis A. Dennery, MD
Scott Lorch, MD
Department of Pediatrics
University of Pennsylvania
School of Medicine
Children’s Hospital of Philadelphia
Philadelphia, PA 19104

REFERENCES

**Paucity of Evidence-Based Research on How to Achieve the Healthy People 2010 Goal of Exclusive Breastfeeding**

To the Editor.—

The evidence base for the public health import of exclusive, prolonged breastfeeding continues to expand, as shown in the article “Breastfeeding and Hospitalization for Diarrheal and Respiratory Infection in the United Kingdom Millennium Cohort Study,” which comes on the heels of the new Healthy People 2010 (HP2010) exclusive breastfeeding goals (objective 16 now reflects exclusive breastfeeding at 16–19 d and e as objectives approved for HP2010). Achieving the new HP2010 goals will require substantial effort, given dismally low rates of exclusive breastfeeding in the United States. Several years ago, *Pediatrics* published an analysis of federal research funding that depicted an incongruity between national priorities for breastfeeding on the one hand and research aimed at achieving Healthy People 2000 (HP2000) breastfeeding goals on the other hand. By the mid-2000s this incongruity had grown starker, even as the new HP2010 targets justifiably raise the bar for this public health objective.

The revised HP2010 breastfeeding objectives adopted in December 2006 set targets for 60% of women to exclusively breastfeed through 3 months and 25% to do so through 6 months. The original HP2010 breastfeeding goals for initiation (75% in early postpartum) and duration (50% and 25% at 3 and 6 months, respectively), did not aim to increase actual intensity of breastfeeding. Rather, these goals for exclusivity were adopted later as part of the HP2010 midcourse review. Although there has been progress toward the original HP2010 objectives, exclusive breastfeeding rates are disturbingly low; just 13% of infants were exclusively breastfed through 6 months, per the American Academy of Pediatrics (AAP) and updated HP2010 guidelines (the AAP statement acknowledged that their Committee on Nutrition supports introducing complementary foods between 4 and 6 months of age when safe and nutritious complementary foods are available). It is notable that in the January/February 2007 issue of *Obstetrics and Gynecology*.
the American College of Obstetrics and Gynecology affirmed the stance of the AAP.3

The revised HP2010 goals follow several federal initiatives. The Surgeon General’s “Blueprint for Action on Breastfeeding” in 2000 led to the National Breastfeeding Awareness Campaign. The campaign featured media outreach, with a specific focus on the risks of not exclusively breastfeeding for 6 months, along with community demonstration projects. Unfortunately, funding for the media campaign ended in 2006. The Maternal and Child Health Bureau does support the AAP’s Breastfeeding Promotion in Physicians’ Office Practices (BPPOP III) program, which has as a major goal the education of health care professionals.6 Although effectiveness measures will be assessed for an evaluation at the project’s completion, BPPOP III was not primarily designed as a research endeavor.

The article “Does Federal Funding for Breastfeeding Research Target Our National Health Objectives?” was published in Pediatrics in 2003.3 The authors identified federally funded research projects relevant to breastfeeding from the years 1994–1996. They reasoned that those projects’ results would be available to inform activities directed toward achieving HP2000 objectives. They searched the CRISP (Computer Retrieval of Information on Scientific Projects) database of federally funded research grants for projects with the keywords “infant nutrition,” “breastfeeding,” or “lactation.”

That analysis identified 362 funded projects from 1994 to 1996 in the area of infant nutrition/breastfeeding/lactation. Of these, just 31 (8.6%) of 362 had as a goal, to increase breastfeeding. Most such grants supported research in the basic science of human milk and lactation. Furthermore, 7.5% of the projects bore no relationship to breastfeeding per se. Rather, they pertained to human milk composition and techniques to improve artificial milks (ie, formula) or develop new pharmaceuticals and therapies. To be sure, the state of existing research around that time was poor. A meta-analysis of published reports from 1996 to 2001 of primary care–based interventions to increase breastfeeding yielded only 22 randomized, controlled trials, just 1 of which was “good quality.”7

Approximately a decade later, there was a sharp drop in the proportion of federally funded research projects aimed at increasing breastfeeding rates in the United States. I conducted a comparable analysis from 2003 to 2006 using the same keywords. The exercise yielded a total of 422 projects. As in the original analysis, titles and abstracts were reviewed to identify projects that had, as a goal, to increase breastfeeding. Of these 422 projects, just 4 specified a goal (direct or indirect) of increasing breastfeeding. Thus, by the mid 2000s, just 4 (<1%) of 422 grants in the area of infant nutrition/breastfeeding/lactation were relevant to the HP2010 objectives for breastfeeding. Basic science continues to dominate.

As clinicians are increasingly expected to practice evidence-based medicine, it is imperative to determine the effects of the above-named strategies and other hospital- and pediatric provider–based practices on breastfeeding duration and exclusivity. Certainly, much remains to be learned about the basic science of human milk. However, unless more infants are actually fed greater amounts of breast milk, for longer periods of time, the knowledge gleaned from this basic science research will have limited utility. At a time when clinical and translational research is being emphasized, a reconsideration of this imbalance in funding priorities in the area of infant feeding would be welcome.

Karen A. Bonuck, PhD
Departments of Epidemiology and Population Health and Obstetrics, Gynecology, and Women’s Health
Montefiore Medical Center/Albert Einstein College of Medicine
Bronx, NY 10467

REFERENCES

doi:10.1542/peds.2007-1210

Much Pain, Little Gain From Voiding Cystourethrogram After Urinary Tract Infection

To the Editor.—

We have read the thoughtful comments on our article1 on the significance of primary vesicoureteral reflux and urinary antibiotic prophylaxis after acute pyelonephritis. Here we want to address the issues of methodologic flaws referred to by Wald.2

much pain, little gain from voiding cystourethrogram after urinary tract infection
AN INTENTION-TO-TREAT ANALYSIS WAS NOT PERFORMED, WEAKENING THE RESULTS OF THE STUDY
To address this concern, we have filled in missing values for each group assuming a worst-case scenario. That is, patients in the nonprophylaxis group who did not complete the study were all assumed to have presented a recurrence of the urinary tract infection and renal scars; patients in the prophylaxis groups who did not complete the study had no urinary tract infection or renal scars. With respect to the recurrence of urinary tract infections (any kind, because we could not assume the localization of the infection), 3 primary questions of interest were considered. First, for patients with vesicoureteral reflux, was there a significant difference in the proportions with recurring infection or renal scars and those with or without prophylaxis? The new analysis showed that there was no significant difference for either infection ($P = .67$) or renal scars ($P = .73$). Second, for patients who did not receive prophylaxis, was there a significant difference in the proportions of those with and without vesicoureteral reflux who experienced infection or renal scars? The reanalysis of the data showed that there was no significant difference in infection ($P = .83$) or renal scars ($P > .99$). Third, was there a significant difference in the proportion of patients who became infected in the vesicoureteral-reflux-with-prophylaxis and the no-vesicoureteral-reflux-without-prophylaxis groups? Again, applying the new analysis of the data, there was no significant difference in infection ($P = .83$) or renal scars ($P = .73$). With this most severe method of analysis, there is still strong support for our primary conclusions.

NEITHER SUBJECTS NOR PHYSICIANS WERE BLINDED TO THE TREATMENT ASSIGNMENTS
We understand that blinding is necessary in certain types of studies in which the Hawthorne effect may play a role. By definition, the Hawthorne effect refers to the improvement seen in patients just because of the special attention paid to the patients during their follow-up. It refers mostly to subjective information. In our opinion, the absence of blinding does not introduce a bias, because the end points in this study (documentation of urinary tract infection and renal scars) were beyond the control of the patients. In addition, the patients who received prophylaxis did not do better than those who did not receive it.

We completely agree that our results should apply only to patients with mild-to-moderate vesicoureteral reflux. We also agree that complete validation of our results will come only when other investigators confirm them.

Eduardo Humberto Garin, MD
Department of Pediatrics
Linda Young, PhD
Department of Biostatistics
University of Florida
Gainesville FL 32610

REFERENCES

doi:10.1542/peds.2007-1233
ERRATA


Errors occurred in the article by Vohra et al, titled “Adverse Events Associated With Pediatric Spinal Manipulation: A Systematic Review,” published in the January 2007 issue of Pediatrics Electronic Pages (doi:10.1542/peds.2006-1392). In the Results section on page e277, paragraph 2, line 5, the authors wrote: “Each case involved a chiropractor and was reported in the United States.” This sentence should be removed. Citation number 37 should be removed from page e277, paragraph 2, lines 10, 15, and 16, and from the Discussion section on page e280, paragraph 1, line 8.

doi:10.1542/peds.2007-1310


Errors occurred in the article by Lee et al, titled “Weight Status in Young Girls and the Onset of Puberty,” published in the March 2007 issue of Pediatrics Electronic Pages (doi:10.1542/peds.2006-2188). In the Methods section on page e625, column 2, under the heading “Puberty Measures,” paragraph 3, lines 7–8, the authors wrote: “Because physical examination did not include breast palpation in this study....” It should read: “Because physical examination did not include breast palpation until 10.5 years of age in this study....” On page e629, column 1, paragraph 2, lines 1–3, the authors wrote: “Neither our study nor the study by Davison et al9 used breast palpation for assessing puberty on physical examination.” It should read: “Similar to the study by Davison et al9, our study did not use breast palpation for assessing puberty on physical examination at 9.5 years of age, although palpation was performed for girls starting at 10.5 years of age.”

doi:10.1542/peds.2007-1351