ARTICLES

Violent Entertainment Pitched to Adolescents: An Analysis of PG-13 Films
Theresa Webb, Lucille Jenkins, Nickolas Browne, Abdelmonen A. Afifi, and Jess Kraus

A Multicenter, Randomized, Double-Blind, Placebo-Controlled Trial to Evaluate the Metabolic and Respiratory Effects of Growth Hormone in Children With Cystic Fibrosis
Dirk Schnabel, Corinna Grasemann, Doris Staab, Hartmut Wollmann, Felix Ratjen for the German Cystic Fibrosis Growth Hormone Study Group

Oral Versus High-Dose Pulse Corticosteroids for Problematic Infantile Hemangiomas: A Randomized, Controlled Trial
Elena Pope, Bernice R. Krafchik, Colin Macarthur, Diana Stempak, Derek Stephens, Miriam Weinstein, Nhung Ho, and Sylvain Baruchel

Propofol Compared With the Morphine, Atropine, and Suxamethonium Regimen as Induction Agents for Neonatal Endotracheal Intubation: A Randomized, Controlled Trial
Satish Ghanta, Mohamed E. Abdel-Latif, Kei Lui, Hari Ravindranathan, John Awad, and Julee Oei

A Randomized, Controlled Trial of a Removable Brace Versus Casting in Children With Low-Risk Ankle Fractures
Kathy Boutis, Andrew R. Willan, Paul Babyn, Unni G. Narayanan, Benjamin Alman, and Suzanne Schuh

Birth Weight, Postnatal Growth, and Risk for High Blood Pressure at 7 Years of Age: Results From the Collaborative Perinatal Project
Anusha H. Hemachandra, Penelope P. Howards, Susan L. Furth, and Mark A. Klebanoff

Firearm Ownership and Storage Patterns Among Families With Children Who Receive Well-Child Care in Pediatric Offices

Chronic Ventilator Need in the Community: A 2005 Pediatric Census of Massachusetts
Robert J. Graham, Eric W. Fleegler, and Walter M. Robinson

To Clean or Not to Clean: Effect on Contamination Rates in Midstream Urine Collections in Toilet-Trained Children
Suzanne Vaillancourt, David McGillivray, Xun Zhang, and Michael S. Kramer
Heelys and Street Gliders Injuries: A New Type of Pediatric Injury
Mihai Vioreanu, Eoin Sheehan, Aaron Glynn, Noelle Casidy, Michael Stephens, and Damian McCormack

Comparison of the Safety and Immunogenicity of a Refrigerator-Stable Versus a Frozen Formulation of ProQuad (Measles, Mumps, Rubella, and Varicella Virus Vaccine Live)
Henry H. Bernstein, Karen Eves, Kristy Campbell, Steven B. Black, Jerry D. Twiggs, Keith S. Reisinger, Ralph M. Conti, Carl-Erik Flodmark, Lars Rombo, Stephanie Klopfner, Florian Schödel, Jonathan Hartzel, Barbara J. Kuter and the Refrigerator-Stable Formulation Study Group for ProQuad

Do Skinfold Measurements Provide Additional Information to Body Mass Index in the Assessment of Body Fatness Among Children and Adolescents?

Perinatal Circulating Visfatin Levels in Intrauterine Growth Restriction
Ariadne Malamitsi-Puchner, Despina D. Briana, Maria Boutsikou, Evangelia Kouskouni, Demetrios Hassiakos, and Demetrios Gourgiotis

Differences in Severity-Adjusted Pediatric Hospitalization Rates Are Associated With Race/Ethnicity
James M. Chamberlain, Jill G. Joseph, Kantilal M. Patel, and Murray M. Pollack

The Spectrum of Cardiac Anomalies in Noonan Syndrome as a Result of Mutations in the PTPN11 Gene
Yves Sznajer, Boris Keren, Clarisse Baumann, Sabrina Pereira, Corinne Alberti, Jacques Elion, Hélène Cavé, and Alain Verloes

Families’ Health-Related Social Problems and Missed Referral Opportunities
Eric W. Fleegler, Tracy A. Lieu, Paul H. Wise, and Sharon Muret-Wagstaff

Visual Function at 11 Years of Age in Preterm-Born Children With and Without Fetal Brain Sparing
Joke H. Kok, Liesbeth Prick, Elly Merckel, Yolande Everhard, Gijs J.Q. Verkerk, and Sicco A. Scherjon

Auditory Brainstem Response Abnormalities and Hearing Loss in Children With Craniosynostosis
Michael W. Church, Leslie Parent-Jenkins, Arlene A. Rozzelle, Frances E. Eldis, and S. Nadya J. Kazzi

One-Year Follow-up of Very Preterm Infants Who Received Lucinactant for Prevention of Respiratory Distress Syndrome: Results From 2 Multicenter Randomized, Controlled Trials
Fernando Moya, Sunil Sinha, Janusz Gadzinowski, Ralph D’Agostino, Robert Segal, Carlos Guardia, Jan Mazela, Genzhiou Liu on behalf of the SELECT and STAR Study Investigators
Use of Sleep Medications in Hospitalized Pediatric Patients
Lisa J. Meltzer, Jodi A. Mindell, Judith A. Owens, and Kelly C. Byars

Compliance With Alarm Limits for Pulse Oximetry in Very Preterm Infants
Luisa Clucas, Lex W. Doyle, Jennifer Dawson, Susan Donath, and Peter G. Davis

A National Outbreak of *Ralstonia mannitolilytica* Associated With Use of a Contaminated Oxygen-Delivery Device Among Pediatric Patients

Temporal Relationships Between Colds, Upper Respiratory Viruses Detected by Polymerase Chain Reaction, and Otitis Media in Young Children Followed Through a Typical Cold Season
Birgit Winther, Cuneyt M. Alper, Ellen M. Mandel, William J. Doyle, and J. Owen Hendley

Immunogenicity and Safety of Intradermal Influenza Immunization at a Reduced Dose in Healthy Children
Susan S. Chiu, J.S. Malik Peiris, Kwok H. Chan, Wilfred Hing Sang Wong, and Yu Lung Lau

Statistical Literacy for Readers of Pediatrics: A Moving Target
Martha A. Hellems, Matthew J. Gurka, and Gregory F. Hayden

Relationship Between Angiotensin-Converting Enzyme Gene Insertion or Deletion Polymorphism and Insulin Sensitivity in Healthy Newborns
Tongyan Han, Xinli Wang, Yunpu Cui, Hongmao Ye, Xiaomei Tong, and Meihua Piao

Infant Swimming Practice, Pulmonary Epithelium Integrity, and the Risk of Allergic and Respiratory Diseases Later in Childhood
Alfred Bernard, Sylviane Carbonnelle, Xavier Dumont, and Marc Nickmilder

Maternal Asthma and Maternal Smoking Are Associated With Increased Risk of Bronchiolitis During Infancy
Kecia N. Carroll, Tebeb Gebretsadik, Marie R. Griffin, William D. Dupont, Edward F. Mitchel, Pingsheng Wu, Rachel Enriquez, and Tina V. Hartert

Laboratory Values for Children With Newly Diagnosed Inflammatory Bowel Disease
David R. Mack, Christine Langton, James Markowitz, Neal LeLeiko, Anne Griffiths, Athos Bousvaros, Jonathan Evans, Subra Kugathasan, Anthony Otley, Mariann Pfefferkorn, Joel Rosh, Adam Mezoff, Susan Moyer, Maria Oliva-Hemker, Robert Rothbaum, Robert Wyllie, J. Fernando delRosario, David Keljo, Trudy Lerer, Jeffrey Hyams for the Pediatric Inflammatory Bowel Disease Collaborative Research Group
Role of Zinc Administration in Prevention of Childhood Diarrhea and Respiratory Illnesses: A Meta-analysis
Rakesh Aggarwal, John Sentz, and Mark A. Miller

Quality of Primary Care and Subsequent Pediatric Emergency Department Utilization
David C. Brousseau, Raymond G. Hoffmann, Ann B. Nattinger, Glenn Flores, Yinghua Zhang, and Marc Gorelick

The Host Hospital 24-Hour Underreferral Rate: An Automated Measure of Call-Center Safety
Daniel A. Hirsh, Harold K. Simon, Robert Massey, Lisa Thornton, and Joseph E. Simon

Intensive Home Visiting Is Associated With Decreased Risk of Infant Death
Edward F. Donovan, Robert T. Ammerman, John Besl, Harry Atherton, Jane C. Khoury, Mekibib Altaye, Frank W. Putnam, and Judith B. Van Ginkel

Endothelial Function in Newborn Infants Is Related to Folate Levels and Birth Weight
Helena Martin, Bo Lindblad, and Mikael Norman

Inhaled Nitric Oxide Therapy Decreases the Risk of Cerebral Palsy in Preterm Infants With Persistent Pulmonary Hypertension of the Newborn
Yuko Tanaka, Tomoshige Hayashi, Hiroyuki Kitajima, Kiyoaki Sumi, and Masanori Fujimura

Patent Ductus Arteriosus and Its Treatment as Risk Factors for Neonatal and Neurodevelopmental Morbidity
Nancy Chorne, Carol Leonard, Robert Piecuch, and Ronald I. Clyman

Continuity of Health Insurance Coverage Among Young Adults With Disabilities
S. Todd Callahan and William O. Cooper

COMMENTARIES

Go Ahead Punk, Make My Day: It's Time for Pediatricians to Take Action Against Media Violence
Victor C. Strasburger

Neonatal Intensive Care Unit Oxygen Management: A Team Effort
Jay P. Goldsmith and Jay S. Greenspan

Drug Pricing in Pediatrics: The Egregious Example of Indomethacin
Alan H. Jobe
Counseling Youth About Military Service Options and Selective Service Registration: An Integral Part of Anticipatory Guidance of Adolescents
Carl E. Stafstrom

Pediatric Prevention of Atherosclerosis: Targeting Early Variation in Vascular Biology
Paul Leeson

Providing a Safety Net for Children: Investigating a Multistate Outbreak of Ralstonia mannitolilytica Related to a Contaminated Reusable Device
Lisa Saiman

Maternal Smoking, Asthma, and Bronchiolitis: Clear-Cut Association or Equivocal Evidence?
Eric A.F. Simões

EXPERIENCE & REASON

Acute Myopericarditis After Multiple Vaccinations in an Adolescent: Case Report and Review of the Literature
Maria T. Thanjan, Prema Ramaswamy, Wyman W. Lai, and Irene D. Lytrivi

Serum Osmolal Gap in Patients With Idiopathic Nephrotic Syndrome and Severe Edema
Gaurav Kapur, Rudolph P. Valentini, Abubakr A. Imam, Amrish Jain, and Tej K. Mattoo

Incidence of Acute Otitis Media and Sinusitis Complicating Upper Respiratory Tract Infection: The Effect of Age
Krystal Revai, Laura A. Dobbs, Sangeeta Nair, Janak A. Patel, James J. Grady, and Tasnee Chonmaitree

The Success of an Immunization Information System in the Wake of Hurricane Katrina
Julie A. Boom, Anna C. Dragsbaek, and Cynthia S. Nelson

FROM THE AMERICAN ACADEMY OF PEDIATRICS

Electronic Prescribing Systems in Pediatrics: The Rationale and Functionality Requirements
Robert S. Gerstle, Christoph U. Lehmann the Council on Clinical Information Technology

Provision of Educationally Related Services for Children and Adolescents With Chronic Diseases and Disabling Conditions
Council on Children With Disabilities

State Children's Health Insurance Program Achievements, Challenges, and Policy Recommendations
Committee on Child Health Financing
Electronic Prescribing Systems in Pediatrics: The Rationale and Functionality Requirements
Council on Clinical Information Technology

Evaluation of Suspected Child Physical Abuse
Nancy D. Kellogg and the Committee on Child Abuse and Neglect

Overuse Injuries, Overtraining, and Burnout in Child and Adolescent Athletes
Joel S. Brenner and the Council on Sports Medicine and Fitness

LETTERS TO THE EDITOR

Score for Neonatal Acute Physiology (SNAP) or Vermont Oxford Risk-Adjustment Model for Very Low Birth Weight Infants?
Luigi Gagliardi and Roberto Bellù

Score for Neonatal Acute Physiology (SNAP) or Vermont Oxford Risk-Adjustment Model for Very Low Birth Weight Infants?: In Reply
John A.F. Zupancic, Shoo K. Lee, and Gabriel J. Escobar

Asthma as a Barrier to Children's Physical Activity
Daniel J. Weiner, Joe McDonough, and Julian Allen

Asthma as a Barrier to Children's Physical Activity: In Reply
Alan Smyth, Ian A. Macdonald, Amy C. McPherson, Judy A. Swift, and Cristine Glazebrook

Subarachnoid Hemorrhage in a Young Child With Sickle Cell Disease: Is Transcranial Doppler Helpful?
Christina S. Ross, Jane M. Hennessy, and Stephen C. Nelson

Subarachnoid Hemorrhage in a Young Child With Sickle Cell Disease: Is Transcranial Doppler Helpful?: In Reply
John J. Strouse, Monica L. Hulbert, Michael R. DeBaun, Lori C. Jordan, and James F. Casella

Neonatal Brain Volumetric Studies: Regression Analysis and Interpretation
George T. Vasileiadis

Neonatal Brain Volumetric Studies: Regression Analysis and Interpretation: In Reply
Nehal A. Parikh, Robert E. Lasky, Kathleen A. Kennedy, and Jon E. Tyson

Early Autism Identification
Penelope J. Pivalizza
Early Autism Identification: In Reply
Vidya B. Gupta, On behalf of the AAP Autism Expert Panel, Paul H. Lipkin On behalf of the AAP Policy Revision Committee on Developmental Surveillance and Screening

Marketing Foods and Beverages: Why Licensed Commercial Characters Should Not Be Used to Sell Healthy Products to Children
Wendi Gosliner and Kristine A. Madsen

Marketing Foods and Beverages: Why Licensed Commercial Characters Should Not Be Used to Sell Healthy Products to Children: In Reply
Victor C. Strasburger

Mycoplasma pneumoniae Encephalitis and Reactivation of Herpes Simplex Virus
Michael Eisenhut

Mycoplasma pneumoniae Encephalitis and Reactivation of Herpes Simplex Virus: In Reply
Ari Bitnun, Jorina M. Elbers, Daune MacGregor, Raymond Tellier, and Susan E. Richardson

Give Synagis Via Home Care: A Reply
Johannes A. Schipper

Give Synagis Via Home Care: A Reply: In Reply
John Goldenring

Pulling the Plug on Entertainment Industry Ratings
Richard Freed

Pulling the Plug on Entertainment Industry Ratings: In Reply
Michael Rich

Adverse Events in the Manipulation of Pediatric Patients: Flaws in a Systematic Review
Anthony Rosner

Adverse Events in the Manipulation of Pediatric Patients: Flaws in a Systematic Review
Murray S. Katz

Adverse Events in the Manipulation of Pediatric Patients: Flaws in a Systematic Review
Richard Brassard
Adverse Events in the Manipulation of Pediatric Patients: Flaws in a Systematic Review: In Reply
Sunita Vohra, Bradley C. Johnston, and Kim Humphreys

Noninitiation or Withdrawal of Intensive Care for High-Risk Newborns
Hugh Whittall

Noninitiation or Withdrawal of Intensive Care for High-Risk Newborns: In Reply
Edward F. Bell, Daniel G. Batton, Ann R. Stark On behalf of the American Academy of Pediatrics Committee on Fetus and Newborn

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

ERRATA

ERRATA

ERRATA

ERRATA

REVIEW ARTICLES

Adherence to Antiretroviral Therapy for Pediatric HIV Infection: A Qualitative Systematic Review With Recommendations for Research and Clinical Management
Jane M. Simoni, Arianna Montgomery, Erin Martin, Michelle New, Penelope A. Demas, and Sohail Rana

SPECIAL ARTICLES

Between Stillness and Story: Lessons of Children's Illness Narratives
Sayantani DasGupta

The Role of Mometasone Furoate Aqueous Nasal Spray in the Treatment of Adenoidal Hypertrophy in the Pediatric Age Group: Preliminary Results of a Prospective, Randomized Study
Marco Berlucchi, Daria Salsi, Luisa Valetti, Giovanni Parrinello, and Piero Nicolai
Violent Entertainment Pitched to Adolescents: An Analysis of PG-13 Films

Theresa Webb, PhD*, Lucille Jenkins, MPH*, Nickolas Browne, EdD*, Abdelmonen A. Afifi, PhD*, Jess Kraus, PhD, MPH*

*Southern California Injury Prevention Research Center, UCLA School of Public Health, and †UCLA School of Film, Television, and Digital Media, University of California–Los Angeles, Los Angeles, California

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

ABSTRACT

OBJECTIVE. The purpose of this study was to evaluate the violence content of the top-grossing PG-13 films of 1999 and 2000 to determine what percentage of it had potential for negative effects on young viewers and what percentage of it had potential for prosocial or beneficial effects.

METHODS. A large, multidimensional analytic instrument was designed for systematic coding of each act of violence and its contextualization by features that have been shown either to enhance or to protect against harmful effects that are associated with violent media exposure: perpetrators and victims of violence, motivation for violence, presence of weapons, degree of realism, and consequences of violence. Descriptive statistics by genre were performed for each film. An ordinal logistic regression model was used to examine the association between the seriousness of violence and weapons, motive, and genre.

RESULTS. In the sample of 77 PG-13 films, a total of 2251 violent actions were observed with roughly half (47%) of lethal magnitude. A total of 118 acts contained justified violence that were initiated by major characters and were extremely serious, and approximately two thirds of the films (49 [64%]) were rated PG-13 for reasons other than violence.

CONCLUSIONS. Violence permeated nearly 90% of the films in our study. Although only a small subset of this content contained violence that was associated with negative effects, only 1 film contained violence that was associated with protective or beneficial effects.
Youth violence is a complex yet commonplace occurrence in American society and involves hundreds of thousands of perpetrators and victims each year. Violent crime statistics for adolescents are indeed staggering. In 2000, 993 youth who were aged ≤18 years were arrested for murder. Between 1980 and 2000, 12- to 17-year-olds committed >35,000 homicides. During these same 2 decades, young people between the ages of 18 and 24 committed an additional 114,797 homicides. Homicide is the second leading cause of death among 15- to 24-year-olds overall and, for more than a decade, the leading cause of death for black individuals.

Causation is without a doubt as complex a matter as the phenomenon of youth violence itself. Research has shown that a whole constellation of interrelated individual, social, and cultural factors must converge for violence to manifest itself in children and youth. The most direct existential method of learning about violence entails witnessing first-hand violent behavior as modeled by family or community members and/or friends. A recent survey of 28,000 sixth-grade public school students in Los Angeles found that the majority of them had been exposed to violence in the previous year.

In addition to direct exposure, a vast and robust body of empirical research shows that exposure to media violence poses a significant risk to the health of children and adolescents. This is predominately the case because media representations of violence model violent behavior and therefore contribute to the learning of violence. This is especially true in contemporary American society, wherein the average young person’s engagement with visual media in all its formats can run to as many as 8 hours a day.

The Hollywood film industry disavows any relationship to education and insists that its only commitment is to entertainment. On this point, the Motion Picture Association of America (MPAA) has been consistent that the goal of Hollywood cinema is temporally to transport and entertain its viewers but in no way to edify or transform them. Social learning theorists, however, have shown that viewers do, in fact, learn from entertainment media. Indeed, popular films can act as powerful teachers engaging children and youths emotionally, even psychologically, in ways that teachers in classrooms can only hope to. Hollywood’s self-appointed watchdog agency, the rating board of the MPAA, was designed and launched in 1968 to stave off government regulation of the film industry. The board is made up of between 8 and 13 nonprofessional raters, including a chairman, who is appointed by the president of the MPAA. It is governed by a code that has 1 articulated standard—“the tastes of the Average American Parent”—and one written rule—“One sexual expletive automatically results in a PG-13 rating. Two sexual expletives automatically result in an R rating, but 1 sexual expletive used in a sexual context is an automatic R.” The stated goal of the MPAA’s rating board is to provide information to parents regarding the content of movies using the following criteria: theme, violence, language, nudity, sensuality, and drug abuse (Fig 1). This board has no executive power, and its recommended ratings can be overturned by an appeals board that is made up of entertainment industry executives.

In addition to the age-based ratings and definitions outlined, the MPAA offers on its Web site as well as in all movie advertisements and packaging on videos and DVDs supplemental content descriptors that broadly characterize the offending material that is contained in each film that it rates more restrictive than G.

---

**FIGURE 1**

The MPAA’s movie rating system.

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>General audiences. All ages admitted.</td>
</tr>
<tr>
<td>PG</td>
<td>Parental guidance suggested. Some material may not be suitable for children.</td>
</tr>
<tr>
<td>PG-13</td>
<td>Parents strongly cautioned. Some material may be inappropriate for children under 13.</td>
</tr>
<tr>
<td>R</td>
<td>Restricted. Under 17 requires accompanying parent or adult guardian.</td>
</tr>
<tr>
<td>NC-17</td>
<td>Under 17 not admitted.</td>
</tr>
</tbody>
</table>

---

The Hollywood film industry disavows any relationship to education and insists that its only commitment is to entertainment. On this point, the Motion Picture Association of America (MPAA) has been consistent that the goal of Hollywood cinema is temporally to transport and entertain its viewers but in no way to edify or transform them. Social learning theorists, however, have shown that viewers do, in fact, learn from entertainment media. Indeed, popular films can act as powerful teachers engaging children and youths emotionally, even psychologically, in ways that teachers in classrooms can only hope to. Hollywood’s self-appointed watchdog agency, the rating board of the MPAA, was designed and launched in 1968 to stave off government regulation of the film industry. The board is made up of between 8 and 13 nonprofessional raters, including a chairman, who is appointed by the president of the MPAA. It is governed by a code that has 1 articulated standard—“the tastes of the Average American Parent”—and one written rule—“One sexual expletive automatically results in a PG-13 rating. Two sexual expletives automatically result in an R rating, but 1 sexual expletive used in a sexual context is an automatic R.” The stated goal of the MPAA’s rating board is to provide information to parents regarding the content of movies using the following criteria: theme, violence, language, nudity, sensuality, and drug abuse (Fig 1). This board has no executive power, and its recommended ratings can be overturned by an appeals board that is made up of entertainment industry executives.

In addition to the age-based ratings and definitions outlined, the MPAA offers on its Web site as well as in all movie advertisements and packaging on videos and DVDs supplemental content descriptors that broadly characterize the offending material that is contained in each film that it rates more restrictive than G.

---

**FIGURE 1**

The MPAA’s movie rating system.

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>General audiences. All ages admitted.</td>
</tr>
<tr>
<td>PG</td>
<td>Parental guidance suggested. Some material may not be suitable for children.</td>
</tr>
<tr>
<td>PG-13</td>
<td>Parents strongly cautioned. Some material may be inappropriate for children under 13.</td>
</tr>
<tr>
<td>R</td>
<td>Restricted. Under 17 requires accompanying parent or adult guardian.</td>
</tr>
<tr>
<td>NC-17</td>
<td>Under 17 not admitted.</td>
</tr>
</tbody>
</table>

---

The Hollywood film industry disavows any relationship to education and insists that its only commitment is to entertainment. On this point, the Motion Picture Association of America (MPAA) has been consistent that the goal of Hollywood cinema is temporally to transport and entertain its viewers but in no way to edify or transform them. Social learning theorists, however, have shown that viewers do, in fact, learn from entertainment media. Indeed, popular films can act as powerful teachers engaging children and youths emotionally, even psychologically, in ways that teachers in classrooms can only hope to. Hollywood’s self-appointed watchdog agency, the rating board of the MPAA, was designed and launched in 1968 to stave off government regulation of the film industry. The board is made up of between 8 and 13 nonprofessional raters, including a chairman, who is appointed by the president of the MPAA. It is governed by a code that has 1 articulated standard—“the tastes of the Average American Parent”—and one written rule—“One sexual expletive automatically results in a PG-13 rating. Two sexual expletives automatically result in an R rating, but 1 sexual expletive used in a sexual context is an automatic R.” The stated goal of the MPAA’s rating board is to provide information to parents regarding the content of movies using the following criteria: theme, violence, language, nudity, sensuality, and drug abuse (Fig 1). This board has no executive power, and its recommended ratings can be overturned by an appeals board that is made up of entertainment industry executives.

In addition to the age-based ratings and definitions outlined, the MPAA offers on its Web site as well as in all movie advertisements and packaging on videos and DVDs supplemental content descriptors that broadly characterize the offending material that is contained in each film that it rates more restrictive than G.

---

**FIGURE 1**

The MPAA’s movie rating system.

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>General audiences. All ages admitted.</td>
</tr>
<tr>
<td>PG</td>
<td>Parental guidance suggested. Some material may not be suitable for children.</td>
</tr>
<tr>
<td>PG-13</td>
<td>Parents strongly cautioned. Some material may be inappropriate for children under 13.</td>
</tr>
<tr>
<td>R</td>
<td>Restricted. Under 17 requires accompanying parent or adult guardian.</td>
</tr>
<tr>
<td>NC-17</td>
<td>Under 17 not admitted.</td>
</tr>
</tbody>
</table>
Our study isolated the PG-13 rating category of films that are targeted at American adolescents. The MPAA’s PG-13 rating category is a gatekeeper and practical construct that is assembled around a disparate set of films that are based on the fundamental warning that defines the category: “Parents strongly cautioned.” Despite that the standards for films that are rated PG-13 include the statement, “Rough or persistent violence is absent,” in recent years, PG-13 has become a repository for action films. These films are often the largest budgeted ones made by the Hollywood film industry and have also been found to be equally, if not more, violent than R-rated films.

The MPAA created the PG-13 rating in 1984 at the behest of Stephen Spielberg, whose PG-rated films Indiana Jones and the Temple of Doom and Gremlins were so saturated with violence that the public was outraged. In 2000, the Federal Trade Commission (FTC) was commissioned by Congress to conduct a study on the marketing practices of the entertainment industries. In an “undercover shopper survey” of 395 theaters across the country, the FTC discovered that unaccompanied children’s access to R-rated fare was successful nearly 50% of the time. As a result, the FTC recommended that movie theater owners begin controlling underage access to the restricted films. In response to this security increase, Hollywood filmmakers began producing more PG-13 films to maximize the potential viewing audience. The box office for the rating category is indicative of this trend. Indeed, the top 10 box office grossing films of 1994–2002 were dominated by the PG-13 rating. During this period, 52% of the films were rated PG-13, followed by R (20%), PG (19%), and G (9%). The PG-13 films grossed 55% ($9.2 billion) of the total box office revenue ($16.7 billion).

Following the format of the National Television Violence Study (NTVS), we examined the instruction in violence that the PG-13 rating category of films offers to youths, its target audience, by focusing not only on the violent act but also on the contextual features that frame the representation of violent actions. We report on the prevalence, seriousness, and graphic nature of screen violence and its contextualization by features that either enhance or protect against the harmful effects that are associated with violent media exposure: perpetrators and victims of violence, motivation for violence, presence of weapons, degree of realism, and consequences of violence.

The underlying justification for reporting on the contextual features of violent action is that it enables us to broaden our perspective on the representation of violence by taking into consideration the larger messages that are generated by the films. In addition, the principal investigators of the NTVS study evaluated the reliability and the validity of each of the previously mentioned measures to ensure that they are grounded theoretically and are “consistent with all existing scientific research assessing the effects of televised violence.”

The questions that we asked were as follows: (1) What percentage of the violence framed by the PG-13 rating category is of the sort that has potential for negative effects on its viewers? (2) Were there any prosocial portrayals of violence in the rating category that would diminish negative effects in viewers or even be beneficial to them? As the juxtaposition of these 2 questions indicates, not all screen violence is problematic. Indeed, some filmic representations of violence produce an antiblackmail message. Last, to evaluate the MPAA rating system’s ability to recognize and identify problematic violence and inform parents thereof, we asked the following: (3) Do the MPAA content descriptors that are attached to the ratings designation systematically account for the violent content that films contained?

**METHODS**

**Study Sample**

The sample included all PG-13–rated films from the 100 top-grossing films of 1999 and 2000 as established by the Hollywood Reporter. The years were chosen because they were the most recent for which complete data were available at the commencement of this study. No previous expectations were made about quantity or level of violent action in the sample. A total of 77 films were included (31 from 1999, 46 from 2000). Film titles are listed in the Appendix. The analytic instrument was modified from an earlier study and used to capture data on violent action and on its corresponding contextual variables. A character index covering demographics, socioeconomic status, and motivations for violence and scales to measure the seriousness and explicitness of the violent action were modified from the instrument that was used in our earlier studies. Questions regarding the presence of weapons and consequences were also used. The genre of each film was abstracted from the MPAA’s Internet-based database.

Two senior researchers from the UCLA Southern California Injury Prevention Research Center served as the coders for all films. The coding used a hierarchical structure that allowed progressive movement from scenes to actions. For each violent action, the film was paused and questions pertaining to the action were answered. Once the coders entered their responses, the film continued until the next violent action. Coders achieved 90% interrater agreement on all variables.

**Definition of Violence**

We defined violence as the representation of any character(s) who intentionally or unintentionally perpetrated physical force or power against another character or group of characters that resulted in (or should have resulted in) physical injury or death. Because we were...
dealing with representational action, we included unintentional or accidental violence (most often situated in visceral comedy wherein a hapless victim falls into a trap intentionally laid by a villain) and counted repetitions of violent acts, thereby broadening the definition of violence from our previous studies.26,31 Our definition corresponds closely to 2 of the NTVS’s 3 “key elements”: emphasis on behavioral acts involving physical force and harmful consequences. In that we do not define “credible threats” as violence is where our study definition differs from theirs.

Measures of Seriousness and Graphic Nature of Violence
To measure the extent or seriousness of violent action, we used a 3-part scale. Level 1 included “pushing, restraining, slapping, pinching, chases without weapons, and spraying or dousing with noxious (but nonlethal) substances.” Level 2 included such behavior as “hitting with a closed fist or with a weapon but with nonlethal force.” Level 3 included all “violent acts executed with deadly force.”

To appraise the degree of explicitness that was used in representing a violent action, we used a 4-part scale. At level 1, “the act was narratively framed but is not itself depicted or, if depicted, with no wounding or expression of serious force.” At level 2, “elements of the act may have been represented, but details of the actual wounding were omitted.” At level 3, “elements of the act may have been represented and the actual wounding may have been shown; however, flesh was not shown to tear, burn, be crushed, or rupture, nor were fragments of body parts represented.” At level 4, “violence was represented in the most explicit manner, and the actual tearing, burning, or otherwise destroying of flesh or body was made visible.”

Presence of Weapons
To capture the presence of weapons involved in the enactment of violence, the instrument provided coders with a list of the following categories: firearms, knives, blunt objects, restraints, explosives, body parts, vehicles, synthetic forces, living creatures, natural forces, and other.

Motivation for Violence
Motivation, or what impels a character to act violently, is the contextual variable that informs the viewer both about why the character acted the way he or she did and whether it was justified. In this study, the motive for each act of violence was selected from 1 of the following 6 categories: protection of life; personal gain; retaliation; to cause pain, injure, or kill; unintentional; and other. Although violent behavior is complex and multiply determined and multiple motivations may be in play in any given act of violence, we selected the dominant and most obvious motive in each case.

Characters
All represented violent actions had a perpetrator and a victim. A perpetrator was the character responsible for provoking an act of violence against another character(s), and the victim was the person who received the physical blow of whatever nature (eg, slap, kick, bullet).

Following the NTVS study, good perpetrators were defined as those who “acted benevolently, helped others, and/or were motivated to consider the needs of others before themselves.” Conversely, the bad perpetrators were egotistical, acted “primarily in their own self-interest, accommodated their own needs, and had very little regard for others.”

For our analysis of perpetrators, we focused only on the major and supporting characters who were actively engaged in violence for the obvious reason that we were able to capture the most information about them. The character analysis was primarily descriptive and weighted according to the number of violent actions in which each character engaged.

Consequences
Scales for bodily, psychological, legal, and medical consequences were developed for this study. The scale for representation of bodily suffering included “no expression,” “the emission of a short expletive or shout without any bodily contortions,” “the emission of a sustained cry of pain without any prolonged bodily contortions,” and “the expression of prolonged or violent bodily contortions from pain with or without the accompaniment of vocal emissions.” The scale that we used to capture the degree of representation of consequences to body included “no representation or only by implication”; “the representation of bruises or slight lacerations to the body”; “the representation of broken bones or serious lacerations”; “bodies maimed, blinded, crippled, scarred for life”; and “dead or dying bodies, fatally wounded, mutilated, decapitated, disememboweled, decomposing.”

Realism
Coders were asked to situate each film in 1 of the following categories: actual reality/true story, fiction, or fantasy.

MPAA Content Descriptors
The MPAA’s Classification and Ratings Administration board places a hierarchy on its supplemental content descriptors (ie, violence, sexuality, language), which provide the rationale and justification for the designated rating. The first descriptor in a series is considered to be the most significant determining element in the rating designation. For each film in the sample, the ranking (first, second, third, etc) of any violence-related factor was recorded. A film that was rated PG-13 for strong language and violence was labeled as having a second-place violence factor. To determine whether the ranking
of the MPAA violence factor reflected the density of violence in each film, we examined the relationship between violence-factor ranking and number of violent acts for each film. The content descriptors were abstracted from the Internet-based MPAA database.

**Statistical Analyses**

Descriptive statistics by genre (e.g., action, comedy) were performed for each film. χ² tests were used to determine the association among genre, MPAA descriptive factors, and frequency of violence. An ordinal logistic regression model was used to examine further the association between the seriousness of the violence and weapon, motive, and genre. Resulting odds ratios (ORs) and P values were used to describe the relationship of seriousness with weapon, motive, and genre. The ordinal logistic regression technique was selected because seriousness, the outcome variable, is a categorical ordered variable with 3 incremental levels. The data were collected and recorded in Filemaker-Pro, managed by Paradox software, and analyzed by using SAS/STAT software.

**RESULTS**

The variables that were examined in this study were derived from our previous research and the NTVS. In the sample of 77 PG-13 films, a total of 2251 violent actions were observed in 67 (87%) of the films (Table 1). Ten films contained no violence, and the remaining 67 films ranged from 1 to 263 violent acts. The overall grand mean number of violent acts in the sample was 29. Seven films ranged from 1 to 263 violent acts. The overall mean number of violent acts in the sample was 29. Seven films had >100 violent acts (Little Nicky; Austin Powers: The Spy Who Shagged Me; The World Is Not Enough; Charlie’s Angels; Crouching Tiger, Hidden Dragon; The Mummy; Mission: Impossible 2; and Shanghai Noon).

In terms of genre, the action films had the highest mean number of violent acts (91 acts per film), followed by the horror/science-fiction films (28 acts per film), comedies (20 acts per film), and drama/romance (4 acts per film). Shanghai Noon, marketed as, “a wildly hilarious, stunt-filled action-adventure comedy,” was the most violent film in the sample (263 acts).

The majority (96%) of the violent acts were intentional. Roughly half (47.9%) of all of the depicted violence in the sample was of level 3 lethal seriousness, or, by definition, “executed with deadly force” (Table 2).

The explicitness of most violent action was level 1 (34.2%) or level 2 (58.7%).

**Presence of Weapons**

Across all genres, the most popular weapon of choice was the body (45.5%), followed by firearms (31.4%; Table 2). Action films had the highest proportion of firearms use (37.7%).

**Motivation for Violence**

Among all films in the sample, the primary motivation for initiating violence was the need to “protect life” (40.6%) followed by the desire “to cause pain, injure, or kill” (27.3%; Table 2).

**Characters**

Of all of the major or supporting characters who initiated violence, 23 were good perpetrators and 11 were bad (Table 3). The good perpetrators initiated 187 acts of violence, and the bad initiated 165 acts. Whereas bad perpetrators were primarily seen in action films (80.0% vs 51.3% good), more good perpetrators than bad appeared in comedies (42.8% vs 11.5%).

The good perpetrators of violence were primarily young adult (20–34 years), white men, whereas the bad perpetrators were more diverse in terms of age, gender, and ethnicity. Approximately two thirds of the bad perpetrators had “efficient, machine-like bodies” and “showed enthusiasm toward violence” compared with only 39.0% and 22.5% of the good perpetrators, respectively. The bad perpetrators never “condemned violence,” whereas 33.2% of the good perpetrators spoke out against violence at some point in the film. Nearly half of the good perpetrators (45.5%) and 32.1% of the bad perpetrators were presented as “cool.”

A total of 299 victims of violence were identified, most of whom were found in action (51.5%) and comedy (32.4%) films. Victims were most often white men and typically not affiliated with any type of violent or criminal-related profession (46.9%).

**Consequences**

Twenty-three bodily consequences and 1 legal and medical consequence were observed in the sample. Among the 23 bodily consequences, 10 (43%) were in action

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Distribution of Violent Actions Among Films According to Genre</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genre</td>
<td>No. (%) of Films with Violence</td>
</tr>
<tr>
<td>Action (n = 13)</td>
<td>13 (16.9)</td>
</tr>
<tr>
<td>Comedy (n = 37)</td>
<td>31 (40.3)</td>
</tr>
<tr>
<td>Drama and romance (n = 18)</td>
<td>14 (18.2)</td>
</tr>
<tr>
<td>Horror, science fiction, thriller (n = 9)</td>
<td>9 (11.7)</td>
</tr>
<tr>
<td>Total (n = 77)</td>
<td>67 (87.0)</td>
</tr>
</tbody>
</table>

* Includes all films in genre, both with and without violence.
films, 9 (39%) were in comedies, 3 (13%) were in drama/romance, and 1 was in horror/science fiction/thriller.

Realism
Twenty-one of the films fell into the category of fantasy, meaning that they contained “characters that could not possibly exist or events that could not possibly happen in the real world as we know it.” This category included such titles as The Mummy; Crouching Tiger, Hidden Dragon; and Battleship Earth. Fifty-six of the films were fictional, with 2 based on true stories. In total, there were 1351 acts of violence in fantasy films (60% of the 2251 total number of acts), 800 (36%) acts in fiction films, and 100 (4%) acts in the 2 films that were based on true stories.

MPAA Content Descriptors
Approximately two thirds of the films (49 films [64%]) did not have an MPAA factor for violence. The mean number of violent acts for these 49 films was 11 (SD: 24; median: 4), with a range of 0 to 107 acts of violence. Of the remaining films, 19 (25%) had a primary MPAA violence factor (mean: 85 acts; SD: 68; range: 3–263 acts; median: 76), and 9 (11%) had a second- through fifth-place violence factor (mean: 11 acts; SD: 10; range: 2–29 acts; median: 3).

As a point of reference, the overall mean number of violent acts in the sample was 29. All action films with >29 acts of violence (n = 9 films) were labeled with an MPAA primary violence factor, yet only one half of the comedies with >29 acts of violence (n = 3 films) were labeled with a primary violence factor. The remaining half of the comedies with >29 acts of violence did not have any corresponding violence factor: The Replacements (32 acts of violence) carried the label, “some crude humor including sexual innuendo and language”; Little Nicky (101 acts) carried the label, “crude sexual humor, some drug content, language and thematic material”; and Austin Powers: The Spy Who Shagged Me (107 acts) stated “sexual innuendo and crude humor.”

Ordinal Logistic Regression
Results from the ordinal logistic regression model summarize the relationship between the act-level variables and seriousness. In 1 unadjusted (crude) analysis, genre was the only covariate. In another analysis, weapon was the only covariate, and in a third analysis, reason for violence was the only covariate. For the weapon variable, responses were collapsed into 3 categories: firearm, body, and other. For the motivation variable, responses were collapsed into 4 categories: protection of life, personal gain, to cause pain, and other. In Table 4, the crude and adjusted ORs are presented for each covariate. The crude ORs do not take into account possible intercorrelations among the covariates, whereas the adjusted ORs do. These results, when juxtaposed against each other,

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Act-Level Characteristics According to Genre</th>
</tr>
</thead>
<tbody>
<tr>
<td>Character</td>
<td>Action (N = 1178)</td>
</tr>
<tr>
<td>----------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Seriousness</td>
<td></td>
</tr>
<tr>
<td>Minimal force (level 1)</td>
<td>81 (6.9)</td>
</tr>
<tr>
<td>Moderate force (level 2)</td>
<td>444 (37.7)</td>
</tr>
<tr>
<td>Lethal force (level 3)</td>
<td>653 (55.4)</td>
</tr>
<tr>
<td>Explicitness</td>
<td></td>
</tr>
<tr>
<td>Level 1</td>
<td>457 (38.8)</td>
</tr>
<tr>
<td>Level 2</td>
<td>666 (56.5)</td>
</tr>
<tr>
<td>Level 3</td>
<td>52 (4.4)</td>
</tr>
<tr>
<td>Level 4</td>
<td>3 (0.3)</td>
</tr>
<tr>
<td>Weapon</td>
<td></td>
</tr>
<tr>
<td>Firearm</td>
<td>444 (37.7)</td>
</tr>
<tr>
<td>Body</td>
<td>459 (39.0)</td>
</tr>
<tr>
<td>Object</td>
<td>181 (15.4)</td>
</tr>
<tr>
<td>Explosion</td>
<td>31 (2.6)</td>
</tr>
<tr>
<td>Vehicle</td>
<td>23 (2.0)</td>
</tr>
<tr>
<td>Other</td>
<td>40 (3.4)</td>
</tr>
<tr>
<td>Reason for violence</td>
<td></td>
</tr>
<tr>
<td>Protection of life</td>
<td>505 (42.9)</td>
</tr>
<tr>
<td>Personal gain</td>
<td>192 (16.3)</td>
</tr>
<tr>
<td>Retaliation</td>
<td>38 (3.2)</td>
</tr>
<tr>
<td>To cause pain, injure, or kill</td>
<td>399 (33.9)</td>
</tr>
<tr>
<td>Other</td>
<td>16 (1.4)</td>
</tr>
<tr>
<td>Unintentional</td>
<td>28 (2.4)</td>
</tr>
</tbody>
</table>

*a P < 0.0001, x² test.
present a comprehensive view of the effects of each variable. The reported ORs are for 1 level of outcome compared with the level just below it. Thus, the OR of 1.33 for action compared with other genres can be interpreted as the ratio of the odds of level 3 seriousness versus level 2 computed for action relative to other genres. The adjusted OR for firearm versus other type of weapon was ~38.

**Synthesis**

The relationship between seriousness, explicitness, repetitive violence, weapon type, and justification with violent acts was examined. An act of violence was considered justified when the primary motivation for violence was to protect life. A total of 118 acts that were initiated by 20 major characters were justified and also extremely serious (level 2 or 3). From this subset of 118 acts, 97% (n = 114) were minimally graphic (level 1 or 2 explicitness), 61% (n = 72) were repeated, and 14% (n = 16) included a firearm as the weapon. Films that contained several of these acts included *The Replacements; Austin Powers: The Spy Who Shagged Me; Mission: Impossible 2; and Battlefield Earth.*

**DISCUSSION**

It is obvious that in the complex sociocultural world of contemporary American society, “no single environ-

---

**TABLE 3**  
Major or Supporting Violent Characters, Act Level

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Bad (n = 165 Acts, 11 Characters)</th>
<th>Good (n = 187 Acts, 23 Characters)</th>
<th>Total (N = 352 Acts) OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genre*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Action</td>
<td>80.0</td>
<td>51.3</td>
<td>64.8</td>
</tr>
<tr>
<td>Comedy</td>
<td>11.5</td>
<td>42.8</td>
<td>28.1</td>
</tr>
<tr>
<td>Drama, romance</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Horror, science fiction, thriller</td>
<td>8.5</td>
<td>5.9</td>
<td>7.1</td>
</tr>
<tr>
<td>Gender*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>41.2</td>
<td>19.8</td>
<td>29.8</td>
</tr>
<tr>
<td>Male</td>
<td>58.8</td>
<td>80.2</td>
<td>70.2</td>
</tr>
<tr>
<td>Apparent age,*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adolescent (13–19)</td>
<td>38.8</td>
<td>0.0</td>
<td>18.2</td>
</tr>
<tr>
<td>Young adult (20–34)</td>
<td>30.9</td>
<td>63.6</td>
<td>48.3</td>
</tr>
<tr>
<td>Adult (35–60)</td>
<td>30.3</td>
<td>36.4</td>
<td>33.5</td>
</tr>
<tr>
<td>Apparent ethnicity*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>52.7</td>
<td>59.9</td>
<td>56.5</td>
</tr>
<tr>
<td>Black</td>
<td>0.0</td>
<td>24.6</td>
<td>13.1</td>
</tr>
<tr>
<td>Asian</td>
<td>38.8</td>
<td>14.4</td>
<td>25.9</td>
</tr>
<tr>
<td>Latino</td>
<td>0.0</td>
<td>1.1</td>
<td>0.6</td>
</tr>
<tr>
<td>Middle Eastern</td>
<td>3.6</td>
<td>0.0</td>
<td>1.7</td>
</tr>
<tr>
<td>NA</td>
<td>4.8</td>
<td>0.0</td>
<td>2.3</td>
</tr>
<tr>
<td>Has efficient, machine-like body*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>68.5</td>
<td>39.0</td>
<td>52.8</td>
</tr>
<tr>
<td>No</td>
<td>31.5</td>
<td>61.0</td>
<td>47.2</td>
</tr>
<tr>
<td>Presented as “cool”*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>32.1</td>
<td>45.5</td>
<td>39.2</td>
</tr>
<tr>
<td>No</td>
<td>67.9</td>
<td>54.5</td>
<td>60.8</td>
</tr>
<tr>
<td>Shows enthusiasm toward violence*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>68.5</td>
<td>22.5</td>
<td>44.0</td>
</tr>
<tr>
<td>No</td>
<td>31.5</td>
<td>77.5</td>
<td>56.0</td>
</tr>
<tr>
<td>Condemns violence*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>0.0</td>
<td>33.2</td>
<td>17.6</td>
</tr>
<tr>
<td>No</td>
<td>100.0</td>
<td>66.8</td>
<td>82.4</td>
</tr>
</tbody>
</table>

*CI indicates confidence interval.  
* Reference category.

**TABLE 4**  
Crude and Adjusted Results From Ordinal Logistic Regression With Seriousness as Dependent Variable

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>n</th>
<th>Crude OR (95% CI)</th>
<th>Adjusted OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genre</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Action</td>
<td>1178</td>
<td>1.33 (0.59–3.02)</td>
<td>0.85 (0.42–1.68)</td>
</tr>
<tr>
<td>Comedy</td>
<td>747</td>
<td>0.65 (0.31–1.33)</td>
<td>0.62 (0.30–1.27)</td>
</tr>
<tr>
<td>Other*</td>
<td>326</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Weapon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firearm</td>
<td>706</td>
<td>4.10 (20.39–82.96)</td>
<td>38.13 (17.53–83.77)</td>
</tr>
<tr>
<td>Body</td>
<td>1025</td>
<td>0.24 (0.14–0.41)</td>
<td>0.24 (0.14–0.43)</td>
</tr>
<tr>
<td>Other*</td>
<td>520</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Reason for violence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protection of life</td>
<td>914</td>
<td>1.39 (0.65–2.97)</td>
<td>1.03 (0.54–1.94)</td>
</tr>
<tr>
<td>Personal gain</td>
<td>362</td>
<td>1.68 (0.73–3.87)</td>
<td>0.94 (0.49–1.79)</td>
</tr>
<tr>
<td>To cause pain, injury, or kill</td>
<td>615</td>
<td>4.91 (2.54–9.49)</td>
<td>2.43 (1.32–4.48)</td>
</tr>
<tr>
<td>Other*</td>
<td>360</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

*CI indicates confidence interval.  
* Reference category.
mental factor can be expected to account for more than a small portion of the individual differences in aggression.”25 However, it is equally obvious that in today’s media-saturated world, education has become indistinguishable from entertainment and that popular films have an impact on beliefs, behaviors, attitudes, and knowledge.31 Media effects research, situated in the paradigm of social learning theory, has fleshed out numerous correlations between media representations of violence and real-world social violence. Each of the 7 contextual variables that we examined in this report embody the potential for a negative or positive impact on viewers. In other words, depending on how they support the violent representation, they can have an inhibitory or a disinhibitory effect, can arouse or discourage aggression, or can facilitate the expression of aggression in the individual viewer.3

**Presence of Weapons**
The “weapons effect,” according to Berkowitz, has to do with the “priming” of aggressive thoughts and behaviors in viewers, which means simply that viewers get ideas from their exposure to media.19,34,35 A multitude of other researchers have found that the presence of weapons, either pictorially or in the natural environment, significantly enhanced aggression among angered as well as nonangered subjects.19,36 Given that >50% of the represented violence in our sample was perpetrated with a weapon, coupled with the fact that most parents are unaware of their potential arousal effect, young viewers of PG-13 films are being subjected unawarely to an enormous amount of problematic violence.

**Motivation for Violence**
The motive for violence provides the viewer the means by which to interpret the action as either justified or unjustified.19 The significance of character motivation was also revealed by Berkowitz, who observed in laboratory experiments that when aroused individuals viewed a scene of violence that they interpreted as justified, they behaved more aggressively than did other individuals in the same condition who interpreted the violence as unjustified.14,37 Most violence in the sample was shown to be justified and was associated with saving or protecting life. Motivations such as self-defense/protection of life and retaliation generally incited the “good” perpetrators to act violently and are accepted by most viewers as reasonable and therefore justified. However, nearly 30% was motivated by a desire to injure another person or to cause pain and suffering. This motivation can hardly be considered justifiable, and unprovoked flights of rage or attacks on innocent targets, the type of violence that are more characteristic of “bad” perpetrators of violence, are usually rejected by viewers as irrational and therefore unjustified.

**Characters**
In terms of characters who are involved in violence—both perpetrators and victims—empirical research has repeatedly shown that viewers of all ages feel a greater affinity toward and generally identify more powerfully with those who are constructed as attractive, likeable, and otherwise “good,” thereby accepting more of what they do as perpetrators and experiencing greater empathy for them as victims. Conversely, characters who are constructed as unattractive, mean, or otherwise antisocial or “bad” are not as appealing to viewers and are therefore less likely to be imitated.5,20 In this study, several of the lead characters were considered “good” yet modeled extremely violent behavior, thereby serving as potential violent role models for viewers.

**Realism**
Realism in the context of fiction has been shown by researchers to have a greater impact on viewers than fantastic portrayals of the same violent action.38 In our sample, 73% of the violence was rendered realistic, but the harm in terms of injury or death that it imposes on its victims was either nonexistent or largely unrealistic, potentially misleading young viewers to believe that violence does not have consequences when in reality it always does. Although a number of effects are associated with this contextual variable, in terms of youth, the learning or imitation effect is the most significant and the 1 of greatest concern to parents and society more generally.

**Consequences**
The vast majority of the violence that was depicted in our sample was inconsequential in terms of bodily expressions of trauma or upset. A consistent practice of elision exists in Hollywood storytelling such that depicted violence is misleading insofar as it pays no attention to the harm that it causes. In light of the fact that studies have shown that the absence of psychological and physical harm in violent portrayals is likely to instigate the learning of aggressive attitudes and behaviors in most viewers, this aesthetic practice is problematic.

Results from experimental research reveal that when consequences to violence are shown, they have an inhibitory effect on the viewer,35 yet only 1 film contained representations of consequential violence that might have a beneficial effect on young viewers. Pay It Forward contained what might be called a prosocial portrayal of violence. At the very least, it contained the sort of violence that discourages imitation. One might argue that it has this effect by establishing audience identification with its young hero and then shattering it when he is stabbed to death by a bully on the school playground at the climax of the film. In addition, the film portrayed the mother’s psychological devastation as well as the community’s reaction to his untimely tragic death.
scenario indeed inspires negative emotions about violence in the viewer. The MPAA described this film as follows: mature thematic elements involving substance abuse/recovery, some sexual situations, language, brief violence. With the exception of this unique film, PG-13 is a frame for what George Gerbner calls “happy violence—cool, swift, and painless,” violence, in other words, without any consideration of its inevitable sequelae: injury and/or death and shattered lives. This type of violence is specifically organized not to upset the viewer. Rather, violence meets “the specifications for the product—conflict visually portrayable, conventions understood by all, attention-drawing action, and [in the context of television] repeated crescendos of suspense amenable to punctuation by commercials.”

MPAA Content Descriptors
The MPAA rating board’s treatment of violence content can be characterized as inconsistent at best. True, the board did relatively well at situating violence as the primary factor when dealing with the most obviously violent genre: action. Indeed, action films for the most part were identified as containing violent content. However, the ratings board failed to identify adequately violent content in the comedies. In terms of the 6 comedic films with >29 violent events, only 3 mention violence in the MPAA content descriptions. Because content descriptors are important for parents who are interested in controlling their children’s exposure to violent content, this oversight on the part of the ratings board is clearly misleading. Even more cause for alarm at this oversight is that comic violence has been shown to be particularly problematic in that it teaches the viewer to associate positive feelings with harming others.

Limitations of the Study
Our sample was not large enough to consider all of the contextual variables that the NTVS analyzed. Furthermore, we examined only 2 years of the top-grossing PG-13 films and are therefore limited by the content of the films that were distributed during that period. However, because there have been no significant changes in the practices of the MPAA rating system, no shift in rating definitions, or any shifts in the film industry, we believe that our sample is representative. In addition, because the study isolated the PG-13 category, no comparison can be made between it and the other ratings categories. Last, because we distinguished violent content and did not examine other factors that are associated with MPAA rating assignment (eg, sexuality, language), our analysis is limited to this subject alone.

Recommendations
First, we believe that the film industry itself needs to reflect on its representational practices and decide which messages it wants to convey to children and adolescents. As it stands, PG-13–rated films include the use of violence as a common means by which conflicts are resolved and stated goals are obtained. We suggest that films that are made for a demographic that is already embroiled in social violence be toned down and that filmmakers begin to make more regular use of the more restricted ratings categories of R and NC-17.

Second, the rating board must become more sensitive to the issue of media violence and acknowledge that the medium that they are responsible for regulating has a learning effect on young viewers. Ample science-based evidence has been produced by social learning theorists documenting the way narrative visual media act as symbolic models and how observation of these models operates as a major mode of acquiring information and knowledge about the social world and which behaviors are appropriate and/or acceptable in that world. Social learning theory should be required reading for MPAA film raters so that familiarity with findings such as that adolescent viewers are most inclined to imitate their screen counterparts, including both the prosocial and the antisocial behaviors modeled by them, would be common knowledge among them.

Third, although parents are often too busy to preview films that their children and adolescents are going to watch and rarely impose restrictions on their media choices as they grow older, we caution them against allowing their children unsupervised viewing of films when violence is identified as the primary or secondary factor in a film or when the word “violence” is substituted with the word “action” and qualified by a generic designation, such as “sci-fi action.” This synonym is indicative generally of high levels of violence played for effect. In addition, numerous Web sites, such as PSVRatings (www.psvratings.com), Kids-in-Mind (www.kidsin-mind.com), and Screen It! (www.screenit.com), provide excellent detailed information about film content.

Fourth, we believe that media literacy education, a discipline that is gaining a foothold in many schools systems across the states, offers a systematic approach to understanding visual media and their messages and the values that they convey. In terms of media violence, McCannon pointed out that a media literacy program that involves parents, limits media diets, and recognizes the value of the media effects research may be promising. Enabling students across the board to engage critically with the essential components of American popular culture using the discipline’s emerging framework would be a major step forward. However, a major challenge for the field of media literacy is that it must have widespread support at the state and national levels, as is done in other countries, to reach its full potentially protective effect.

Finally, we strongly recommend that pediatricians and public health professionals continue to play an advocacy role both locally with educators and school...
boards and nationally with policy-makers to support the development of a more child-friendly media environment. We wish to reinforce the American Academy of Pediatrics recommendation that pediatricians also take a leadership role in keeping media literacy on the public health agenda and where possible work with entertainment industry officials to develop a content-based ratings system to help parents make healthy media choices. With unified involvement of the community, parents, pediatricians, researchers, educators, caregivers, media producers, and the government, the quest to mitigate the effects of media violence is achievable.

CONCLUSIONS
The findings from this study align with those of the NTVS, which evaluated >10 000 hours of television programming from 1994 to 1997. Violence permeated almost 90% of the films in our study and 90% of movies that were viewed for the NTVS. Weapons (other than the body itself) commonly appear in both film and television: >50% of the acts of film violence and 26% of the NTVS programming. In both studies, much of the violence was inconsequential and situated within a humorous context, which is troubling when one considers that this type of media violence is deemed problematic in terms of triggering aggressive behavior in viewers.

ACKNOWLEDGMENTS
Support for this research was provided by the Southern California Injury Prevention Research Center, which is funded by US Centers for Disease Control and Prevention grant 5 R49 CE00199-02.

We thank the referees and associate editor for helpful comments.

REFERENCES
6. Landsberg M. Majority of L.A. 6th graders see violence: either as victims or witnesses, students are often traumatized and schoolwork suffers. Los Angeles Times. 2005;B California 6
22. Waxman S. Rated S for secret. Washington Post. April 8, 2001;G1
24. Motion Picture Association of America. What do the ratings mean? Available at: www.mpaa.org/FilmRatings.asp. Accessed November 1, 2006


46. Thoman E, Jolls T. Literacy for the 21st Century: An Overview and Orientation Guide to Media Literacy Education. Santa Monica, CA: Center for Media Literacy; 2005

---

**Appendix:** Film List

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>The Sixth Sense</td>
<td>Cast Away</td>
</tr>
<tr>
<td>Austin Powers: The Spy Who Shagged Me</td>
<td>Mission: Impossible 2</td>
</tr>
<tr>
<td>Big Daddy</td>
<td>What Women Want</td>
</tr>
<tr>
<td>The Mummy</td>
<td>The Perfect Storm</td>
</tr>
<tr>
<td>The World Is Not Enough</td>
<td>Meet the Parents</td>
</tr>
<tr>
<td>Notting Hill</td>
<td>X-Men</td>
</tr>
<tr>
<td>Wild Wild West</td>
<td>What Lies Beneath</td>
</tr>
<tr>
<td>The Haunting</td>
<td>Crouching Tiger, Hidden Dragon</td>
</tr>
<tr>
<td>Entrapment</td>
<td>Charlie’s Angels</td>
</tr>
<tr>
<td>Blue Streak</td>
<td>Nutty Professor II: The Klumps</td>
</tr>
<tr>
<td>Boxwinger</td>
<td>Big Momma’s House</td>
</tr>
<tr>
<td>She’s All That</td>
<td>Miss Congeniality</td>
</tr>
<tr>
<td>The Cider House Rules</td>
<td>Gone in 60 Seconds</td>
</tr>
<tr>
<td>Never Been Kissed</td>
<td>Unbreakable</td>
</tr>
<tr>
<td>Forces of Nature</td>
<td>Space Cowboys</td>
</tr>
<tr>
<td>Message in a Bottle</td>
<td>U-571</td>
</tr>
<tr>
<td>Anna and the King</td>
<td>The Family Man</td>
</tr>
<tr>
<td>10 Things I Hate About You</td>
<td>Chocolat</td>
</tr>
<tr>
<td>For Love of the Game</td>
<td>Vertical Limit</td>
</tr>
<tr>
<td>Mickey Blue Eyes</td>
<td>Bring It On</td>
</tr>
<tr>
<td>Superstar</td>
<td>Mission to Mars</td>
</tr>
<tr>
<td>Mystery Men</td>
<td>Coyote Ugly</td>
</tr>
<tr>
<td>The Out-of-Towners</td>
<td>Shanghai Noon</td>
</tr>
<tr>
<td>The Other Sister</td>
<td>Finding Forrester</td>
</tr>
<tr>
<td>Blast From the Past</td>
<td>Dude, Where’s My Car?</td>
</tr>
<tr>
<td>EDtv</td>
<td>Frequency</td>
</tr>
<tr>
<td>At First Sight</td>
<td>The Replacements</td>
</tr>
<tr>
<td>Anywhere but Here</td>
<td>O Brother, Where Art Thou?</td>
</tr>
<tr>
<td>An Ideal Husband</td>
<td>Little Nicky</td>
</tr>
<tr>
<td>Drive Me Crazy</td>
<td>Bedazzled</td>
</tr>
<tr>
<td>A Midsummer Night’s Dream</td>
<td>Autumn in New York</td>
</tr>
<tr>
<td>28 Days</td>
<td>Keeping the Faith</td>
</tr>
<tr>
<td>Keeping the Faith</td>
<td>Bounce</td>
</tr>
<tr>
<td>Hanging up</td>
<td>Bring It On</td>
</tr>
<tr>
<td>The Skulls</td>
<td>The Replacements</td>
</tr>
<tr>
<td>Thirteen Days</td>
<td>O Brother, Where Art Thou?</td>
</tr>
<tr>
<td>The 6th Day</td>
<td>Little Nicky</td>
</tr>
<tr>
<td>Where the Heart Is</td>
<td>Pay It Forward</td>
</tr>
<tr>
<td>Pay It Forward</td>
<td>The Legend of Bagger Vance</td>
</tr>
<tr>
<td>The Skulls</td>
<td>Love and Basketball</td>
</tr>
<tr>
<td>Thirteen Days</td>
<td>Boys and Girls</td>
</tr>
<tr>
<td>The 6th Day</td>
<td>Battlefield Earth</td>
</tr>
<tr>
<td>Where the Heart Is</td>
<td>Down to You</td>
</tr>
<tr>
<td>Pay It Forward</td>
<td>Best in Show</td>
</tr>
</tbody>
</table>
A Multicenter, Randomized, Double-Blind, Placebo-Controlled Trial to Evaluate the Metabolic and Respiratory Effects of Growth Hormone in Children With Cystic Fibrosis

Dirk Schnabel, MD, Corinna Grasemann, MD, Doris Staab, MD, Hartmut Wollmann, MD, PhD, Felix Ratjen, MD, PhD, for the German Cystic Fibrosis Growth Hormone Study Group

Departments of Pediatric Endocrinology and Diabetology and Pediatric Allergy and Pulmonology, Children’s Hospital, Charite, Berlin, Germany; Children’s Hospital, University of Duisburg, Essen, Germany; Pfizer Endocrine Care, New York, New York; University Children’s Hospital, Tubingen, Germany; Division of Respiratory Medicine, Hospital for Sick Children, Toronto, Ontario, Canada

Financial Disclosure: This study was supported by Pharmacia. Dr. Wollmann is a former employee of Pharmacia.

ABSTRACT

OBJECTIVE. Positive effects of growth hormone therapy on growth, nutritional status, and lung function have been observed in patients with cystic fibrosis, but the current evidence is based on unblinded studies that involved a small number of patients. This trial was designed as a multicenter, randomized, placebo-controlled, double-blind study to assess the efficacy and safety of 2 dosages of growth hormone in cystic fibrosis.

METHODS. Sixty-three dystrophic patients with cystic fibrosis were randomly assigned for 24 weeks to 1 of 3 treatment arms: growth hormone dosage of 0.11 IU/kg body weight per day, growth hormone dosage of 0.21 IU/kg body weight per day, or placebo. The 24-week double-blind period was followed by an open treatment period of 24 weeks. The primary outcome measure was the change in forced expiratory volume in 1 second in percentage predicted from baseline. Secondary outcome measures were changes in height, weight, and exercise tolerance.

RESULTS. Height, growth velocity, and growth factors (insulin-like growth factor 1 and insulin-like growth factor–binding protein 3) increased significantly in both treatment groups, whereas weight gain did not differ between the growth hormone groups and placebo. A trend toward improvement in absolute forced vital capacity was observed in patients who received the higher growth hormone dosage, whereas forced expiratory volume in 1 second did not change significantly with growth hormone treatment. Maximal oxygen uptake during peak exercise increased significantly in treated patients. There were no significant differences in the frequency or severity of adverse events or in the incidence of abnormalities in glucose metabolism.

CONCLUSIONS. These data suggest that in the group investigated, growth hormone therapy was well tolerated and had positive metabolic effects but did not result in short-term improvement of lung function in patients with cystic fibrosis.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2783
doi:10.1542/peds.2006-2783

Key Words
cystic fibrosis, growth hormone, clinical trial, lung function

Abbreviations
CF—cystic fibrosis
CFTR—cystic fibrosis transmembrane regulator
GH—growth hormone
rhGH—recombinant human growth hormone
BIA—bioelectrical impedance analysis
IGF—insulin-like growth factor
IGFBP—insulin-like growth factor–binding protein
FEV1—forced expiratory volume in 1 second
V̇O2—oxygen uptake
ITT—intention to treat
PP—per protocol
AE—adverse event
SAE—severe adverse event
Accepted for publication Dec 12, 2006
Address correspondence to Felix Ratjen, MD, PhD, Division of Respiratory Medicine, Hospital for Sick Children, 555 University Ave, Toronto, Ontario, Canada M5G 1X8. E-mail: felix.ratjen@sickkids.ca
Cystic Fibrosis (CF) is caused by defects in the cystic fibrosis transmembrane regulator (CFTR) protein, which functions as a cyclic adenosine monophosphate-regulated chloride channel. The CFTR defect results in a multisystem disorder with the dominant clinical features being chronic lung disease and exocrine pancreatic insufficiency. Advances in medical treatment, including intensive antibiotic treatment and aggressive nutritional support, have resulted in improved prognosis with increasing numbers of patients surviving into adult life.\(^1\)\(^2\) However, normal growth and body weight are not achieved in all patients.\(^3\) Catabolism has been recognized as a poor prognostic marker in these patients, because it is correlated with reduced life expectancy.\(^4\)\(^5\)\(^6\)

Malnutrition is a major cause of growth retardation and results from loss of exocrine pancreatic function, increased energy requirements, and systemic anorexia in patients with severe pulmonary disease manifestation. Chronically increased concentrations of inflammatory cytokines such as tumor necrosis factor α and interleukin 1 and 6 affect the production and the secretion of growth hormone (GH) or growth factors but can also inhibit growth at the tissue level as a result of GH resistance.\(^7\)\(^8\) Overall, patients with CF have been shown to be shorter than predicted by their genetic target range.\(^3\)\(^5\)

Recombinant human GH (rhGH) as an anabolic adjunct to stimulate growth has been studied in a number of trials that reported benefits on growth, nutritional status, and lung volume.\(^2\)\(^7\)\(^-\)\(^22\) All published trials were conducted unblinded and involved a small number of patients. This study was undertaken as the first double-blind, placebo-controlled trial to determine whether the addition of GH to standard therapy has beneficial metabolic and pulmonary effects in patients with CF.

**METHODS**

The trial was designed as a multicenter, randomized, double-blind, placebo-controlled, parallel-groups study to compare the efficacy and the safety of 2 fixed dosages of rhGH or placebo in patients with CF. The study was conducted as a 24-week double-blind study with a 24-week open-label treatment period in 12 German CF centers between January 2001 and February 2004. The study was approved by the institutional review boards of the participating centers. Written informed consent was obtained in all cases.\(^4\) Patients who were included in this trial had an established CF diagnosis that was based on a sweat chloride concentration >60 mmol/L and/or 2 disease-causing CFTR mutations, a bone age of 8 to 18 years, dystrophy defined as a BMI <10th and/or body weight <3rd percentile despite a high caloric intake (>120% of the recommended dietary allowance) according to a 3-day food-intake diary. Exclusion criteria were an acute pulmonary exacerbation in the 4 weeks before entering the trial; diabetes (fasting plasma glucose > 126 mg/dL); liver cirrhosis with hypoalbuminemia; serum creatinine > 120 μmol/L; inability to perform exercise and lung-function testing; history of malignancy; suspected non-compliance; participation in any other clinical trial during active treatment phase; pregnancy or lactation; and treatment with GH, anabolic steroids, or systemic corticosteroids within 12 months before the start of the study. Eligible patients underwent a screening visit with clinical assessment and lung-function testing. The patient’s body composition was determined by means of bioelectrical impedance analysis (BIA) measured by BIA 2000-M (Data Input, Frankfurt, Germany). Fat mass, lean body mass, and extracellular and total body water were recorded. Radiograph of the patient’s left wrist was performed for bone age according to Greulich and Pyle.\(^23\) Patients were also instructed to complete a 3-day food-intake history. Patients who fulfilled the entry criteria were entered into the trial 1 week after the screening visit and reassessed at 12-week intervals thereafter.

Standing height was recorded as the means of 3 independent measurements by use of a wall-mounted Harpenden stadiometer ( Holtain Ltd, Crymych, Wales). Body weight was measured on a standardized scale (Soehnle S 20, Murrhardt, Germany). Height and body weight values were compared with the standards of Reinken and van Oost.\(^24\) BMI was assessed by the standards of Rolland-Cachera et al.\(^25\)

Patients were assigned to 1 of the 3 treatment arms for 24 weeks in a double-blind manner with 0.070 mg/kg body weight per day (≈0.21 IU/kg body weight per day) somatropin (subsequently referred to as higher dosage), 0.039 mg/kg body weight per day (≈0.11 IU/kg body weight per day) somatropin (subsequently referred to as lower dosage), or placebo. Recombinant somatropin from *Escherichia coli* K12 (Genotropin; Pfizer, Karlsruhe, Germany) was supplied in 2-chamber cartridges with powder and solvent for injection using a Genotropin Pen. For safety reasons, the maximum daily GH dosage was limited to 3.4 mg. After the end of the double-blind treatment period, patients who were on GH therapy were maintained on their current GH dosage for an additional 24 weeks; patients in the placebo group were randomly assigned to either the low or the high GH-dosage treatment regimen.

Assessments at the 12-week intervals included pulmonary-function measurements with a spirometer as well as blood sampling to measure growth factors (insulin-like growth factor 1 [IGF-1] and IGF-binding protein 3 [IGF-BP3]). In addition glucose, cholesterol, triglycerides, electrolytes, creatinine, serum urea nitrogen, albumin, liver function test (aspartate aminotransferase, alanine aminotransferase, alkaline phosphatase, and bilirubin), and immunologic parameters (red blood cells, white blood cells, thrombocytes, and immunoglobulins) were determined at these time points. An oral glucose tolerance test was performed at study entry as well as
every 24 weeks during treatment. The patient received 1.75 g/kg body weight (maximum 75 g) of oral glucose, and plasma levels were determined immediately before glucose intake as well as after 120 minutes.

At the 24-week intervals, exercise testing using a bicycle ergometer was performed. Initial work rate was set at 20 W and increased every minute by increments of 15 W for patients who were between 125 and 150 cm or 20 W for patients who were taller than 150 cm. Ten-watt increments were used for patients with a forced expiratory volume in 1 second (FEV₁) percentage predicted of 30% or less. Heart rate, O₂ consumption (V̇O₂), CO₂ production, and ventilatory volumes were measured continuously while the patient was breathing through a mouthpiece. The exercise test was terminated at subjective exhaustion. A Borg scale was used to assess subjective exertion at each exercise testing.

A patient diary card that contained information on fever >38.5°C, increased cough or sputum, intake of additional and/or other antibiotics, days off school/work, and hospitalization was dispensed/collected at each visit. Quality of life was assessed every 12 weeks using a disease-specific questionnaire. Clinical safety variables (physical assessment, systolic and diastolic arterial blood pressure, and heart rate) were determined at each study visit.

The study was conducted according to a predefined analysis plan using a 3-stage group sequential adaptive design with sample size adjustments after planned interim analyses. For the purpose of efficacy analysis, 2 data sets were to be defined: the intention-to-treat (ITT) sample and the per-protocol (PP) sample. The ITT sample included all patients who had received at least 1 dose of study medication and who had been assessed at least once after baseline. The PP sample, being a subset of the ITT sample, included patients who did not violate the study protocol or terminate treatment prematurely.

The primary end point was the mean change from baseline in FEV₁ at the end of the 24-week double-blind treatment period. Secondary objectives were changes in exercise capacity, body weight (lean body mass), lung function, hospitalization, days off school or work, and quality of life.

At the first interim analysis, the 2 null hypotheses H₀₁ and H₀₂ were to be rejected and the study stopped if the analysis for FEV₁ yielded a P value of <.00026. If H₀₁ or both H₀₁ and H₀₂ could not be rejected, then the study was to be continued with a recalculated sample size based on the effect size estimation of the interim analysis. At the second confirmatory analysis, null hypotheses that could not yet be rejected at the first interim analysis were to be rejected if the test statistic that is based on the inverse normal method exceeded the critical value 2.454.

For the hypotheses that could not be rejected at the second stage, the sample size was recalculated again and the analogous procedure was conducted at the third (final) analysis by using the critical value 2.004. This procedure preserved the overall type I error rate of α = .025. For confirmatory hypothesis testing at the interim analysis as well as at the final analysis, the nonparametric Mann-Whitney U and Wilcoxon rank sum W tests (1-sided) were used. All other group comparisons are hypothesis generating in nature (ie, P values that resulted from statistical tests were interpreted as exploratory). The statistical analysis was conducted by ClinResearch GmbH (Cologne, Germany) using SPSS (SPSS, Chicago, IL) and SAS (SAS Institute, Cary, NC) software packages.

**RESULTS**

The study was powered to show a significant effect of GH on FEV₁. The first interim analysis was planned with 12 patients in each treatment arm and conducted with 14 patients in the low-dosage group, 12 patients in the high-dosage group, and 10 patients in the placebo group (36 patients according to ITT). The second interim analysis was planned with 24 patients in each treatment arm and conducted with 20 patients in the low-dosage group, 20 patients in the high-dosage group, and 19 patients in the placebo group (59 patients according to ITT). The second interim analysis revealed that a total number of 300 patients would be needed to achieve significance for FEV₁. Because recruitment of 300 patients for this study was not feasible in the participating centers, additional recruitment was stopped at this time point.

Seventy-one patients were available for the final statistical analysis. Four patients refused study participation before randomization. Sixty-four of the 67 randomly assigned patients who took study medication at least once were contained in the safety data set. The primary analysis of efficacy according to the ITT analysis was based on 63 patients (24 female, 39 male) who were randomly assigned, took study medication, and were assessed at least once after baseline. The baseline characteristics for the ITT study population are summarized in Table 1. In addition, a PP analysis of a reduced set of variables was performed on 51 patients (19 lower dosage, 16 higher dosage, and 16 placebo).

**Lung Function**

FEV₁ was primarily analyzed as change from baseline in SD scores. Compared with baseline values, FEV₁ did not change significantly during the 6-month period (Table 2). In addition, no difference in percentage change from baseline FEV₁ was observed in GH-treated compared with placebo-treated patients with CF (P = .362; Fig 1). There was no correlation between baseline lung function and changes in FEV₁ during the study period (data not shown). Although a trend toward an increase in absolute forced vital capacity was seen both in the lower- and higher-dosage groups (Table 2, Fig 2), this difference was not statistically significant compared with placebo for
Changes in FEV1 in percentage from baseline during the 24-week double-blind study period in the higher-dosage (■), lower-dosage (□), and placebo (○) groups. Data reflect mean ± SEM for the 3 groups. No significant differences were observed among the groups.

the low-dosage group ($P = .49$) and failed to reach statistical significance for the high-dosage group ($P = .068$). Flows at lower lung volumes (maximal expiratory flow at 50% of vital capacity and maximal midexpiratory flow) showed no systematic changes during the study period (data not shown). No major changes were observed in FEV1 during the open treatment phase in any of the groups (mean change ± SD: 0.6 ± 2.4% for all groups). The mean increase in forced vital capacity in the 24 weeks of open trial was 3.1 ± 5.5% with no significant differences between the groups.

**Growth and Weight**

Absolute growth velocity increased significantly in both the higher- ($P = .025$) and the lower-dosage groups ($P = .018; \text{Fig } 3, \text{ Table 2}$). In parallel, growth velocity in SD scores for chronological age increased in both treated groups. The mean increase in forced vital capacity in the 24 weeks of open trial was 3.1 ± 5.5% with no significant differences between the groups.

### TABLE 1
Baseline Characteristics of the Study Population

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Higher Dosage (n = 20)</th>
<th>Lower Dosage (n = 22)</th>
<th>Placebo (n = 21)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean ± SD (range), y</td>
<td>14.3 ± 2.6 (9.0 to 18.0)</td>
<td>13.8 ± 2.7 (9.0 to 19.9)</td>
<td>14.6 ± 2.9 (11.0 to 20.0)</td>
</tr>
<tr>
<td>Bone age, mean ± SD (range), y</td>
<td>13.1 ± 2.5 (13 to 18)</td>
<td>12.3 ± 2.0 (8 to 15)</td>
<td>12.6 ± 2.4 (9 to 17)</td>
</tr>
<tr>
<td>Height, mean ± SD (range), cm</td>
<td>151.7 ± 12.5 (120 to 168)</td>
<td>151.3 ± 10.5 (124 to 171)</td>
<td>149.8 ± 11.7 (134 to 176)</td>
</tr>
<tr>
<td>Height, mean ± SD (range), SDS</td>
<td>−2.1 ± 1.1 (−3.6 to 0.3)</td>
<td>−1.8 ± 1.3 (−4.1 to 0.9)</td>
<td>−2.5 ± 1.2 (−4.6 to 0.6)</td>
</tr>
<tr>
<td>Weight, mean ± SD (range), kg</td>
<td>36.5 ± 7.7 (19.3 to 49.7)</td>
<td>35.4 ± 7.5 (21.6 to 51.2)</td>
<td>34.6 ± 6.7 (24.4 to 48.4)</td>
</tr>
<tr>
<td>BMI, mean ± SD (range), SDS</td>
<td>−2.0 ± 1.0 (−4.6 to 0.0)</td>
<td>−2.0 ± 1.0 (−4.1 to 0.9)</td>
<td>−2.2 ± 0.8 (−4.1 to 0.6)</td>
</tr>
<tr>
<td>Fat mass, mean ± SD, kg</td>
<td>5.32 ± 2.43</td>
<td>4.04 ± 2.46</td>
<td>4.59 ± 2.75</td>
</tr>
<tr>
<td>FEV1, mean ± SD, SDS</td>
<td>−1.14 ± 0.77</td>
<td>−1.17 ± 0.66</td>
<td>−1.03 ± 0.65</td>
</tr>
<tr>
<td>FVC, mean ± SD, % predicted</td>
<td>66.0 ± 15.4</td>
<td>67.5 ± 15.1</td>
<td>66.6 ± 15.3</td>
</tr>
<tr>
<td>FEV1, mean ± SD, % predicted</td>
<td>51.7 ± 20.1</td>
<td>54.2 ± 21.6</td>
<td>55.1 ± 19.1</td>
</tr>
<tr>
<td>MEF25-75%VC, mean ± SD, % predicted</td>
<td>32.6 ± 24.9</td>
<td>37.6 ± 33.9</td>
<td>36.1 ± 24.4</td>
</tr>
<tr>
<td>IGF-1, mean ± SD (range), ng/mL</td>
<td>233 (113 to 616)</td>
<td>256 (75 to 1000)</td>
<td>206 (59 to 328)</td>
</tr>
<tr>
<td>IGFBP-3, mean ± SD (range), mg/mL</td>
<td>4.6 (2.5 to 6.8)</td>
<td>4.29 (2.34 to 11.4)</td>
<td>3.83 (1.72 to 6.03)</td>
</tr>
</tbody>
</table>

SDS indicates SD score; FVC, forced vital capacity; MEF25-75%VC, maximal midexpiratory flow.

### TABLE 2
Changes in Metabolic and Respiratory Parameters During the 24-Week Double-Blind Study Period

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Higher Dosage</th>
<th>Lower Dosage</th>
<th>Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth velocity, cm/y</td>
<td>6.8 ± 4.3a</td>
<td>5.6 ± 2.9a</td>
<td>3.5 ± 2.3</td>
</tr>
<tr>
<td>Growth velocity, SDS</td>
<td>2.6 ± 2.7a</td>
<td>1.5 ± 2.6a</td>
<td>−1.0 ± 7.2</td>
</tr>
<tr>
<td>Weight increase, kg</td>
<td>2.2 ± 2.3</td>
<td>2.4 ± 1.9</td>
<td>1.4 ± 1.7</td>
</tr>
<tr>
<td>Lean body mass, kg</td>
<td>2.3 ± 2.5</td>
<td>2.5 ± 2.4</td>
<td>1.5 ± 2.3</td>
</tr>
<tr>
<td>Fat mass, kg</td>
<td>−0.75 ± 1.27</td>
<td>−0.45 ± 2.07</td>
<td>0.03 ± 2.3</td>
</tr>
<tr>
<td>BMI, SDS</td>
<td>0.1 ± 0.6</td>
<td>0 ± 0.6</td>
<td>0.1 ± 0.4</td>
</tr>
<tr>
<td>IGF-1, SDS</td>
<td>−0.04 ± 0.3</td>
<td>−0.03 ± 0.32</td>
<td>−0.03 ± 0.44</td>
</tr>
<tr>
<td>FEV1, %</td>
<td>4.3 ± 13.4</td>
<td>3.5 ± 12.3</td>
<td>1.0 ± 23.0</td>
</tr>
<tr>
<td>FVC, %</td>
<td>6.0 ± 11.2</td>
<td>3.1 ± 13.1</td>
<td>−0.7 ± 15.1</td>
</tr>
<tr>
<td>Exercise test work rate, W</td>
<td>6.0 ± 36.2</td>
<td>11.4 ± 18.0</td>
<td>1.6 ± 17.8</td>
</tr>
<tr>
<td>$\bar{V}O_2$, ml/min</td>
<td>26.4 ± 77.2a</td>
<td>12.5 ± 28.5a</td>
<td>2.4 ± 17.0</td>
</tr>
<tr>
<td>IGF-1, SDS</td>
<td>1.66 ± 1.53a</td>
<td>1.00 ± 1.06a</td>
<td>−0.37 ± 1.24</td>
</tr>
<tr>
<td>IGFBP-3, SDS</td>
<td>0.66 ± 1.41a</td>
<td>0.5 ± 1.6a</td>
<td>−0.15 ± 0.78</td>
</tr>
</tbody>
</table>

All parameters are shown as changes from baseline. SDS indicates SD score.

$a P \leq .05$ versus the placebo group.
significantly different between treatment groups (Table 2, Fig 5). Changes in weight in percentage from baseline were negatively correlated with age ($r = 0.32; P = .05$) and positively correlated with baseline lung function ($r = 0.49; P = .001$) (Fig 6). Similar relationships were not observed in the placebo group. Body composition reflected by body fat and lean body weight was very similar at baseline in the 3 groups (Table 1). Lean body mass increased in all 3 groups. Fat mass remained unchanged in the placebo group and decreased marginally in the 2 treatment groups (Table 2). None of these differences between groups was statistically significant.

IGF-1 and IGFBP-3

IGF-1, the metabolic mediator of anabolic GH action and growth-stimulating factor, increased in treated patients and remained unchanged in the placebo group ($P < .002$ for both treatment groups; Table 2). In parallel, IGFBP-3, the main binding protein of IGF-1, increased, whereas concentrations decreased in the placebo group ($P < .05$ for treated patients).

Exercise Tolerance

At baseline, there were no significant differences between the treatment groups in work rate, peak heart rate, maximal VO$_2$, CO$_2$ production, O$_2$ saturation, minute volume, tidal volume, and subjective exertion as assessed by the Borg scale at the end of the exercise test. A nonsignificant increase in work rate was observed in both treatment groups, whereas work rate was essentially unchanged in the placebo group (Table 2). Subjective exertion as monitored on the Borg scale was similar in the various treatment groups; there were no signifi-
cant differences in O₂ saturation both at baseline and during peak exercise between the groups.

Absolute Vo₂ increased in both the lower- and higher-dosage groups and remained unchanged in the placebo group (P = .05 versus placebo; Fig 7). These differences were more pronounced in the PP analysis (P = .009 for the lower-dosage group and P = .02 for the higher-dosage group). During the open treatment period, both mean work rate (6.1 ± 16.6 W) and mean Vo₂ at peak exercise (86.9 ± 220.4 mL/min) increased in patients who previously received placebo.

The assessment of the quality-of-life questionnaire revealed no major differences among the treatment groups. There were no statistically significant differences in the number of days off school or work as well as the number of hospitalization days among the 3 groups.

Compliance
Compliance was assessed by diaries and vial counts returned at regular study visits. According to this analysis, 19 (95.0%) of 20 patients in the higher-dosage group, all 22 (100.0%) patients in the lower-dosage group, and 19 (90.5%) of 21 patients in the placebo group confirmed the regularly received study medication.

Safety Measures
Twelve (60.0%) of 20, 14 (63.6%) of 22, and 13 (59.1%) of 22 patients in the higher-dosage, lower-dosage, and placebo groups, respectively, experienced at least 1 adverse event (AE) during the double-blind study period. The most frequently observed single category was pulmonary exacerbations, affecting 7 (35.0%) of 20, 6 (27.3%) of 22, and 4 (18.2%) of 22 patients in the higher-dosage, lower-dosage, and placebo groups, respectively, with no significant differences among the groups.

All treatment groups were approximately equally affected by severe AEs (SAEs), namely 4 (20.0%) of 20, 5 (22.7%) of 22, and 4 (18.2%) of 22 patients in the higher-dosage, lower-dosage, and placebo groups, respectively. The distribution of SAEs was similar to that of nonserious AEs, with respiratory infections requiring antibiotic therapy being the most frequently reported SAE (Table 3).

Fasting blood glucose levels (high dosage: 91.9 ± 7.0 mg/dL; low dosage: 95.9 ± 12.0 mg/dL; placebo: 87.6 ± 10.1 mg/dL) as well as stimulated glucose levels (high dosage: 117.0 ± 62.2 mg/dL; low dosage: 109.9 ± 31.6 mg/dL; placebo: 108.0 ± 37.9 mg/dL) after oral glucose tolerance testing were almost identical at baseline. No significant changes in fasting glucose (high dosage: 96.8 ± 13.4 mg/dL; low dosage: 101.2 ± 15.2 mg/dL; placebo: 88.8 ± 13.7 mg/dL) and stimulated glucose concentrations (high dosage: 120.7 ± 59.5 mg/dL; low dosage: 128.2 ± 39.5 mg/dL; placebo: 116.1 ± 23.3 mg/dL) were seen after the 24-week double-blind, placebo-controlled period. GH treatment did not result in an inappropriate acceleration of bone age (data not shown).

**DISCUSSION**
To our knowledge, this is the first study to assess the effect of GH therapy in patients with CF in a double-blind, placebo-controlled design. GH treatment increased growth velocity and improved maximal Vo₂ during peak exercise but had no significant effect on pulmonary function. Therefore, GH demonstrated positive metabolic effects but did not result in short-term improvement in pulmonary status.

The increase in height and growth velocity that was observed in this study is in concordance with previous unblinded studies.9–12 Younger children grew better in response to GH therapy than did older patients. This study therefore provides additional evidence that linear growth rate can be increased with GH therapy in patients with CF. Nevertheless, it can be debated whether the dosages used in this study are sufficient to optimize growth in this patient population. IGF-1 and IGFBP-3

<table>
<thead>
<tr>
<th>SAE</th>
<th>Higher Dosage</th>
<th>Lower Dosage</th>
<th>Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulmonary exacerbation</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Hemoptysis</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Pneumothorax</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Productive cough</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Candida sepsis</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Distal intestinal obstruction syndrome</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hyperglycemia</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Convulsion</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Port-a-cath blockage</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Ligament rupture</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

**FIGURE 7**
Changes in maximal Vo₂ (Vo₂ max) during peak exercise in percentage from baseline during the 24-week double-blind study period in the higher-dosage (■), lower-dosage (□), and placebo (○) groups. Data reflect mean ± SEM for the 3 groups. There was a significant increase in Vo₂ max in both treatment groups (P = .05).
Levels were normalized in both groups but did not exceed the reference ranges. This study provides no clear evidence for a dosage effect, because no significant differences were observed in any of the outcome variables between the 2 treatment groups. Dosages ranging from 0.16 to 0.35 mg/kg body weight per week have been used in previous trials, and the higher-dosage group in this study therefore received the highest dosage that has been given to patients with CF to date. This dosage is also higher compared with studies that were performed in other populations without primary GH deficiency. Although it cannot be excluded that a more pronounced metabolic effect can be achieved with even higher dosages, this study provides no evidence to support this hypothesis.

Weight gain was observed in all groups in this study, but no significant differences were detected between GH-treated and placebo-treated patients. The mean yearly increase in weight was 4.4 and 5 kg in the 2 previous randomized trials, which is similar to the overall weight gain that was observed during this trial. It is interesting that placebo-treated patients in this study also demonstrated a similar weight gain, which may reflect a nonspecific effect driven by study participation and more intense nutritional counseling. This underlines the need to assess treatment effects in a blinded study to clarify whether changes are directly related to therapy. Although no significant difference was observed in body weight, a trend toward higher lean body mass and lower fat mass was observed in treated patients and likely reflects the anabolic effects of GH. Whether shifts in body mass are associated with improvements in clinical status in CF is unclear. In addition, techniques that assess body composition, such as BIA, have not been thoroughly validated as an outcome parameter for clinical trials in the CF population.

A previous study described impressive weight gains after GH treatment in patients who had CF and received nutritional support via enteral tube feeding. Enteral tubes are often placed if weight gain is insufficient in patients with CF. Enteral tube feeding has been demonstrated to benefit most but not all of the patients with nutritional failure. Only 3 patients who were included in this study (all in the lower-dosage group) received enteral tube feeding as part of their routine medical care. This may reflect a selection bias for the patients who were included in this study, because enteral tube feeding is usually offered to patients who have CF with inadequate caloric intake. Because of the low number of patients who received enteral tube feeding, we could not perform a meaningful subgroup analysis. The inclusion criteria of this trial required patients to have a high caloric intake independent of the way that these calories were administered. As reflected by the previous study that focused on enteral tube-fed patients alone, weight gain may be more pronounced in these patients. Whether a high-calorie diet that is provided for extended periods, as is the case for enteral tube feeding, will improve the anabolic potency of GH could not be addressed by this study.

Pulmonary function in percentage predicted did not change significantly with GH therapy in patients with CF, although a trend toward an increase in absolute lung volumes was seen in the higher-dosage group, which supports the findings of previous unblinded trials. Whether an increase in lung volume that parallels an increase in body size can be considered a positive finding is debatable. Most of the previous treatment approaches that directly targeted the lung did not use growth as an outcome parameter, and no long-term data are available to clarify whether increases in height and absolute lung function will improve the subsequent course of the disease process. Because the primary effects of GH in CF are metabolic rather than pulmonary, the potential respiratory benefits of GH treatment may not translate into short-term increases in lung function. In addition, the lack of short-term decline in FEV, seen in recent trials in patients with CF, including this study, have made it difficult to use lung function as a primary outcome parameter. Whether GH therapy has a stabilizing effect on lung function over time is much harder to assess and would require a longer treatment and observation period. Long-term studies to assess treatment effects on evolution of lung function over time have become challenging, with overall improvements in CF therapy that have resulted in an annual decline in pulmonary function of <2%.

Positive effects of GH therapy on work load and maximal Vo2 during exercise were observed in this trial. Maximal Vo2 during exercise has been reported to predict mortality in patients with CF and significant lung disease, and improvements in exercise capability may increase physical activity levels and eventually patients' outcome. An inactive lifestyle with lower physical activity has been proposed to be partially responsible for the more rapid decline of female patients during puberty. Whether an intervention such as GH therapy will positively affect physical activity and whether this would result in better outcome for patients with CF cannot be answered with this study design.

Compared with previous studies, this trial included older patients with CF and more advanced pulmonary disease. The mean FEV1 in these patients is considerably lower than the mean FEV1 of the overall CF population at a similar age, and this may reflect the well-described link between nutritional failure and poor lung function. The positive correlation between baseline pulmonary function and weight gain would suggest that patients with poorer lung function are unlikely to benefit from GH therapy. In addition, benefits on growth were more pertinent in younger patients. Therefore, treatment can
be expected to be more efficacious if started in younger patients with preserved pulmonary function.

No SAEs of GH that could be directly attributed to GH therapy were observed in this trial. Because GH may increase blood glucose level, concerns had been raised that GH therapy would trigger diabetes in patients with CF. Random blood glucose levels or glycosylated hemoglobin concentrations are not sufficient to diagnose abnormal glucose metabolism in CF, and only glucose tolerance testing can be considered adequate to address this question.44 Similar to previous trials, none of the patients developed glucose impairment as investigated by oral glucose tolerance test during the study. This trial confirms the evidence from previous trials that GH therapy seems to be safe in children with CF with regard to glucose metabolism.

CONCLUSIONS

This short-term study provides evidence for positive metabolic effects of GH therapy in dystrophic patients with CF. Although growth velocity and exercise capacity as well as biochemical marker of anabolism improved with treatment, no direct effects on pulmonary function could be observed. Long-term studies in a larger cohort of patients are needed to clarify whether GH therapy will have a positive effect on lung-function decline in patients with CF. In addition, IGF-1 therapy may be a plausible alternative to GH that warrants additional study. This study suggests that the effect of GH treatment may be better if less severely affected and younger patients are included. As an end point, a combination of morphologic (eg, body composition), functional (eg, exercise capacity), pulmonary (eg, FEV$_1$), and biochemical markers might be useful to assess the efficacy of treatment.

ACKNOWLEDGMENTS

Investigators for the German Cystic Fibrosis Growth Hormone Study were Michael Barker (Aachen), Dirk Schnabel (Berlin, principal investigator), Doris Staab (Berlin), Volker Stephan (Bochum), Ernst Rietschel (Koeln), Antje Schuster (Düsseldorf), Lutz Nährlich (Erlangen), Felix Ratjen (Essen), Hans-Georg Posselt (Frankfurt), Joachim Kühr (Freiburg), Sabine Brömme (Halle/Saale), Manfred Ballmann (Hannover), and Matthias Griese (München).

REFERENCES

Oral Versus High-Dose Pulse Corticosteroids for Problematic Infantile Hemangiomas: A Randomized, Controlled Trial

Elena Pope, MD, MSc, Bernice R. Krafchik, MB, ChB, Colin Macarthur, MD, PhD, Diana Stempak, PhD, Derek Stephens, MSc, Miriam Weinstein, MD, Nhung Ho, MD, Sylvain Baruchel, MD

Section of Dermatology, New Agents and Innovative Therapy Program, and Research Institute, Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada; Bloorview Research Institute, Bloorview Sick Kids Rehabilitation, Toronto, Ontario, Canada

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

ABSTRACT

OBJECTIVES. Oral systemic corticosteroids are the mainstay of treatment for problematic hemangiomas; however, current information is based on anecdotal experience and retrospective studies. We aimed to determine whether systemic steroids are efficacious in proliferating hemangioma and to compare the efficacy and safety of 2 corticosteroid treatment modalities.

PATIENTS AND METHODS. Twenty patients with problematic hemangiomas of infancy were randomly assigned to either daily oral prednisolone or monthly intravenous pulses of methylprednisolone. Their clinical outcomes (improvement using a visual analog score) and adverse events were compared at 3 months from baseline and 1 year of age. Data on possible surrogate markers of angiogenesis were available for the first 3 months.

RESULTS. At 3 months, orally treated patients had a median visual analog score of 70 compared with 12 in the intravenous group. This response pattern was similar at the patients’ first birthday: 50.0 vs 11.5. Additional treatment beyond 3 months was needed for 65% of the patients (7 in the intravenous and 6 in the oral group). Six of 8 patients with impaired vision at enrollment had an improved function at 1 year (4 patients in the intravenous group and 3 patients in the oral group). Of the 4 surrogate markers of angiogenesis measured (plasma basic fibroblast growth factor, vascular endothelial growth factor, vascular cellular adhesion molecule 1, and endoglin, and urine basic fibroblast growth factor), the only 2 that decreased over time were vascular cellular adhesion molecule 1 and endoglin. Patients in the oral group had a higher rate of adverse effects, such as hypertension (18.6% vs 13.1%), abnormal cortisol (78% vs 60%), and growth retardation.

CONCLUSIONS. Systemic corticosteroids are efficacious in stopping the proliferation of hemangiomas. The oral corticosteroids offered more clinical and biological benefit than the pulse steroids with higher risk of adverse effects.
HEMANGIOMAS ARE THE MOST COMMON BENIGN TUMORS OF INFANCY, OCCURRING IN ~10% OF CHILDREN BY 1 YEAR OF AGE. These lesions are seen more frequently in females, premature infants, and twins. Endothelial proliferation is the hallmark of this condition. Proangiogenic factors, such as vascular endothelial growth factor (VEGF) and cellular adhesion molecules (intercellular adhesion molecule 3, E-selectin, and VCAM-1), are increased in the proliferative phase of hemangiomas, whereas basic fibroblast growth factor (bFGF) is higher in the resolution phases. It is unknown, however, how and when these angiogenesis factors are turned off. It is also unclear whether predictions of clinical behavior and/or therapeutic decisions could be made on the basis of the circulating levels of the angiogenesis factors.

Although most hemangiomas resolve spontaneously, 10% to 20% require treatment because of interference with function and/or significant disfigurement. Current guidelines recommend treatment for (1) life- and function-threatening hemangiomas (vision impairment, airway obstruction, congestive heart failure, and hepatic involvement), (2) large, disfiguring facial hemangiomas, (3) hemangiomas in locations that may lead to permanent scarring or deformity (eg, nose, ear, etc), and (4) ulcerated hemangiomas. Systemic corticosteroids administered orally are, at present, the mainstream of treatment for problematic hemangiomas. The mechanism of action of steroids is unclear but seems to be related to inhibition of angiogenesis. Information regarding dose, length of therapy, and weaning schedule for steroids is based on anecdotal experience and retrospective studies. Most children are treated with doses of 2 to 4 mg/kg per day for 4 to 12 weeks, followed by slow weaning over several months. This regimen is associated with a 30% to 84% response rate but leads to with various degrees of adverse events. Pulse steroids, supraphysiological doses of glucocorticoids, are usually used in conditions where rapid effects are desired and are associated with manageable, transient, short-term complications. Pulse steroids, alone or in combination with other modalities, have been used in diffuse hemangiomatosis and in other vascular tumors, such as hemangioendotheliomas.

The objective of this study was to determine whether high-dose intravenous pulse corticosteroids are more efficacious in reducing the size of hemangiomas and safer than oral corticosteroids. A secondary objective was to measure changes in the putative surrogate markers of angiogenesis over time.

PATIENTS AND METHODS

Study Population and Study Design

The study was conducted between July 2002 and July 2005 at the Hospital for Sick Children (SickKids), an academic pediatric tertiary referral center. The study received ethical approval from the institution’s research ethics board and has been registered at www.clinicaltrials.gov (NCT 00312520).

Eligible infants were between 1 and 4 months of age and had “problematic” facial infantile hemangiomas (IHs), defined as periorbital/orbital tumors with visual impairment and/or large size/disfiguring hemangiomas. Infants >4 months of age, those with concomitant congenital heart disease, and those with nonfacial IHs were excluded.

The families and the principal investigator (PI) who followed the patients were not blinded to the intervention. However, the assessors who measured the primary outcome were blinded to the patient’s intervention allocation.

Intervention

The study had 2 treatment groups. In the oral group, infants received oral prednisolone, 2 mg/kg per day, in 2 divided doses for 3 months. This dose was followed by a tapering schedule (decreasing the dose by 1 mg per month) over 6 to 9 months to prevent rebound. The IV group received pulses of intravenous high-dose corticosteroids monthly for 3 months. A pulse consisted of methylprednisolone in doses of 30 mg/kg per day infused over 1 hour daily for 3 days. The study design allowed for infants to receive additional treatment beyond the 3 months if there was evidence of rebound or ongoing proliferation. For the oral group it meant retreatment with a second course of oral corticosteroids, whereas for the IV group it involved either monthly pulse steroids if, in the opinion of the PI, the infant had response but needed additional treatment or oral corticosteroids if there was no response with the IV pulse steroids (treatment failure).

The study design allowed for these patients to have oral steroids added to their regimen if there was significant rebound or worsening of the lesion between pulses (such patients were considered treatment failures for the final analysis). Patients were offered continuation of the pulses if they responded but required additional treatment beyond the 3 months. Patients in the oral group were prescribed concomitant oral ranitidine to minimize steroid-related gastrointestinal adverse effects.

Subjects were allocated randomly to each group by the research pharmacist who prepared blocks of 4. The PI monitored all of the subjects every 2 weeks for the first month then monthly for 6 months and every 2 months thereafter. The study had 2 major end time points: 3 months from the enrollment (time point 1) and at the subjects’ first birthday (time point 2).

Outcomes

The primary outcome was change in the size of the hemangioma. Serial photographs (front and side view) were taken using a standardized approach (background,
distance, and angle from the patient). Assessors (blinded to patient allocation) and parents were asked to compare photographs and rate change in the hemangioma at 3 months versus baseline and at 1 year versus baseline, using a 100-mm visual analog scale (VAS) with a range of −100 to +100, where “0” represented no change, “+” represented a decrease in size and “−” represented an increase in the size of the hemangioma.

Secondary outcomes included (1) change in the visual function at 1 year in infants with periorbital hemangiomas, (2) adverse events captured using parent diaries (behavior changes, irritability, crying, hyperactivity, apathy, insomnia, vomiting, and abdominal pains), medical charts (blood pressure, heart rate, and respiratory rate), and investigations (complete blood cell [CBC] count, blood sugar, renal function tests, electrolytes, and morning cortisol), and (3) changes over time in angiogenesis markers. All of the blood pressure (BP) readings were done manually with size-appropriate cuffs. If an abnormal reading was obtained, second and third readings, 15 minutes apart, were performed. The lowest value of the 3 readings was recorded. Hypertension was defined as persistent BP readings >95th percentile for the patient’s age.

The angiogenesis markers tested were plasma bFGF, VEGF, VCAM-1, endoglin, and urine bFGF at baseline and 1, 2, and 3 months. All of the blood and urine samples were placed on ice and centrifuged at 2000 rpm for 10 minutes at 4°C within 15 minutes. The plasma and sediment-free urine was aliquoted and frozen at −80°C until analyses using the following commercially available enzyme-linked immunosorbent assay kits, Quantikine human VEGF, bFGF, VCAM-1, and endoglin immunoassay kits (R&D Systems, Inc, Minneapolis, MN), according to the manufacturer’s directions. Plasma results were expressed in picograms per milliliter for VEGF and bFGF and in nanograms per milliliter for VCAM-1 and endoglin. Urine bFGF results were expressed as picograms per gram of creatinine.

**Statistical Analysis**

A total of 20 patients (10 in each group) provided a power of 80% (2-tailed α level of .05) to detect a difference between the 2 groups of ≥20% in their VAS scores. Descriptive statistics included means, medians, and proportions. The intraclass correlation coefficient was used to assess the reliability of assessments. For the primary outcome (VAS), the nonparametric Mann-Whitney U test was used to test the differences between groups. For the secondary outcomes, continuous variables were tested using the Mann-Whitney U test, whereas differences between groups on categorical variables were tested using Fisher’s exact test. Repeated-measures analyses were used to test differences between the 2 groups on the surrogate angiogenesis markers. A 2-sided P of .05 indicated statistical significance. Statistical analyses were conducted using SAS 9.1 (SAS Institute Inc, Cary, NC).

**RESULTS**

**Characteristics of the Study Population**

Twenty patients, 10 in each study group, met the inclusion and exclusion criteria. The mean age at enrollment was 12 (±4) weeks, with female predominance (85%). Hemangiomas were present at birth in 45% of patients. All of the patients had mixed, superficial and deep facial hemangiomas, with an equal segmental distribution (Table 1). There were 2 patients with prematurity, 1 of them having a birth weight of <1500 g. Four patients had a family history of hemangiomas. Two patients, 1 in each group, presented with ulcerations. The 2 groups were not statistically different in their baseline characteristics (Table 1); however, the IV group had more subjects with periorbital involvement at enrollment (5 vs 3). Imaging to assess the depth of the lesions and associated abnormalities was performed in 7 patients (3 patients had MRIs, 1 patient had computed tomography, and 1 patient had ultrasonography). One patient had

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Oral Group</th>
<th>IV Group</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at start (SD), wk</td>
<td>11 (4)</td>
<td>12 (3)</td>
<td>.55</td>
</tr>
<tr>
<td>Female gender, n/N</td>
<td>8/10</td>
<td>9/10</td>
<td>1.00</td>
</tr>
<tr>
<td>Hemangioma at birth, n/N</td>
<td>5/10</td>
<td>4/10</td>
<td>1.00</td>
</tr>
<tr>
<td>Ulceration, n/N</td>
<td>0/10</td>
<td>2/10</td>
<td>.47</td>
</tr>
<tr>
<td>Facial segments, n/N</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1-frontotemporal</td>
<td>1/10</td>
<td>4/10</td>
<td>.39</td>
</tr>
<tr>
<td>S2-maxillary</td>
<td>3/10</td>
<td>1/10</td>
<td></td>
</tr>
<tr>
<td>S3-mandibular</td>
<td>3/10</td>
<td>3/10</td>
<td></td>
</tr>
<tr>
<td>S4-frontonasal</td>
<td>3/10</td>
<td>2/10</td>
<td></td>
</tr>
<tr>
<td>Eye involvement, n/N</td>
<td>3/10</td>
<td>5/10</td>
<td>.17</td>
</tr>
<tr>
<td>Family history, n/N</td>
<td>2/10</td>
<td>2/10</td>
<td>1.00</td>
</tr>
<tr>
<td>Birth weight (SD), kg</td>
<td>3.1 (0.9)</td>
<td>3.3 (0.59)</td>
<td>.83</td>
</tr>
<tr>
<td>Weight at baseline (SD), kg</td>
<td>5.4 (1)</td>
<td>5.9 (0.9)</td>
<td>.26</td>
</tr>
<tr>
<td>Height at baseline (SD), cm</td>
<td>57 (6)</td>
<td>59 (3)</td>
<td>.32</td>
</tr>
</tbody>
</table>
evidence of prominent extra-axial cerebrospinal fluid spaces related to prematurity. All of the patients completed the study.

Efficacy

Clinical Improvement Using VAS
At both time points, the infants in the oral group had greater improvement in the size of the hemangiomas compared with the IV group. For example, at time point 1 (3 months compared with baseline) the median VAS score was 70 (interquartile range [IQR]: 54 to 80) in the oral group compared with a median VAS of 12 (IQR: −18 to 39) in the IV group (P = .002). At time point 2 (1 year of life compared with baseline), the median VAS was 50 (IQR: 35 to 67) in the oral group versus a median of −1.5 (IQR: −35 to 22) in the IV group (P = .005). There was good agreement between the 2 blinded assessors (intraclass correlation coefficient: 0.6; 95% confidence interval: 0.21–0.81). Similarly, there was good correlation between assessors’ and parents’ VAS scores (0.92). At 3 months, patients with periorbital hemangiomas had less improvement compared with other areas of facial involvement, with a median VAS of 11 (IQR: −19 to 55) versus a median of 61 (IQR: 23 to 78; P = .049). Similar findings were noted at 12 months, with a median VAS of 4 (IQR: −24 to 36) versus a VAS of 48 (IQR: 9 to 72; P = .049).

Functional Improvement
Forty percent (8 of 20) of the patients had evidence of eye involvement at enrollment as determined by a pediatric ophthalmologist (3 in the oral group and 5 in the IV group). This consisted of amblyopia in 6 (75%) of 8 patients, astigmatism in 3 (38%) of 8, and increased intraocular pressure in 5 (63%) of 8. Two patients had no change in their eye findings at 1 year, 1 in each treatment group. The other patients (6 of 8) had improvement in their eye findings (4 patients in the IV group and 2 patients in the oral group), suggesting that despite the lack of significant changes in the appearance of the lesions (as assessed by VAS), treatment was efficacious.

Need for Additional Treatment
Thirteen patients required additional treatment (65%). Seven patients (54%) belonged to the IV group, and 6 patients (46%) were in the oral group. In the oral group, the additional treatment consisted of a second course of 2 mg/kg per day of corticosteroids for 4 to 6 weeks during the weaning phase and was because of regrowth of the lesion. For the IV group, 6 patients required an average of 2.6 additional pulses (range: 1–6) until 1 year of age. In addition, 2 patients received intralesional steroids after their last pulse. One patient in the IV group required surgery because of the inability of the corticosteroids to significantly decrease the size of the lesion.

Surrogate Markers of Angiogenesis
Four angiogenesis markers, plasma bFGF, VEGF, VCAM-1, and endoglin, as well as urine bFGF, were measured at baseline and every month for the first 3 months of the study in a subgroup of 13 patients (7 in the oral group and 6 in the IV group). The angiogenesis markers that decreased significantly over time were VCAM-1 (P < .001) and endoglin (P = .03), whereas the others did not change over the study period (Fig 1). Moreover, a statistical difference between the oral group and IV group values for VCAM-1 and endoglin was found in 2 of 3 follow-up visits, mirroring the clinical regression of the hemangiomas (Fig 2).

Safety

Parental Reports
There was no difference between the 2 groups with respect to irritability, excessive crying, apathy, insomnia, vomiting, abdominal pain, or behavioral changes (Table 2). One patient in the oral group had persistent vomiting over several weeks, which required changing the formulation of prednisolone to dose-equivalent dexamethasone.

Physician-Documented Adverse Events
There were no adverse events noted during pulse steroid infusions. In the oral group, 72% of BP measurements were ≤90th percentile (31 of 43), 9.4% were 75th to 95th percentile (4 of 43), and 18.6% were ≥95th percentile (8 of 43). In the IV group, 76% of BP measurements were ≤90th percentile (35 of 46), 10.9% were 90th to 95th percentile, and 13.1% were ≥95th percentile. The patients with measurements ≥95th percentile were asked to have their BP monitored by their regular doctor on a weekly basis. Only 1 patient (in the oral group) required antihypertensive medication for persistent high BP. Two patients experienced serious adverse events (respiratory distress) requiring hospital admission, 1 in each group. Both patients had uneventful recoveries. The patient in the oral group also developed uncomplicated chickenpox infection. There was no difference between the patients’ growth parameters at 3 months (P = .13 for weight and P = .3 for height). However, infants in the oral group had evidence of growth retardation at 1 year of age as documented by the differences in their weight (P = .003) and height (P < .001) (Fig 3).

Investigations
Renal function tests, serum electrolytes, CBC count, morning cortisol, and blood sugar were measured in all of the patients at baseline and monthly for 3 months. There were no clinically significant abnormalities noted
FIGURE 1
Entire study-population plots of angiogenic markers over time.
in the renal function tests, electrolytes, and CBC counts. Fifty-two plasma morning cortisol tests \( (n = 73 \% 73) \) were found to be abnormal, 31 in the oral group \( (n = 38 \% 82) \) and 21 in the IV group \( (n = 35 \% 60) \). Twelve cortisol levels in the oral group and 1 in the IV group were in the undetectable range, suggestive of significant adrenal suppression. Blood sugar was found to be mildly elevated in 2 of 38 tests in the oral group and 3 of 32 tests in the IV group. In all of the patients with an abnormal blood sugar, the abnormality was transient.

**DISCUSSION**

Approximately 10% to 20% of infants with hemangiomas require intervention to prevent or decrease complications related to their proliferation. Systemic corticosteroids are the most common treatment modality used in the management of problematic hemangiomas. This is the first randomized, controlled trial to demonstrate the efficacy of corticosteroids in halting the proliferation of IHs clinically (decrease in hemangioma size and improved visual function) and biologically (decrease in the circulating proangiogenic proteins). The effect, measured over the first year of life, was more pronounced with continuous, daily doses of prednisolone than with high-dose, intermittent pulses with methylprednisolone. One of the biggest limitations in designing studies in IHs is quantifying the response to treatment. To date, there are no good clinical, laboratory, or imaging tools that can accurately size these lesions. We adapted a validated tool (VAS) to estimate size differences between visits by comparing digital photographs. However, medical photography has its own limitations (the inability to quantify the
volume of the lesion, the difficulty in standardizing the photographs, and the difficulty in accurately distinguishing changes because of the natural growth of the face from changes related to proliferation).

Our efficacy data are comparable to previous reports. Several case series have shown that 30% of patients have a significant improvement after steroid initiation, 30% show no change requiring additional dose increases to 5 mg/kg per day, and 40% have an equivocal response.12,17 In a recent meta-analysis,15 using a mean steroid dose of 2.9 mg/kg per day, 84% of patients had cessation of growth or regression of the hemangiomas, and the rebound (increase in the size of the mass after an initial shrinkage) rate was 36% (95% confidence interval: 29%–44%). Several retrospective studies suggested that higher doses are associated with a higher clinical response.14,15 In a study using high-dose oral methylprednisolone (30 mg/kg per day for 5 days with tapering every 5 days for a total of 6 weeks), a high initial response rate with high doses was noted; however, the overall response seemed to be similar to 5 mg/kg per day.26 Given the lower rate of adverse effects and shorter duration of treatment, this regimen was preferable to longer, lower-dose courses.26 In our study, the superiority of higher doses could not be duplicated (median VAS of 12 vs 70 at time point 1 and 1.5 vs 50 at time point 2). Another surprising finding was the fact that VAS scores were lower at 1 year compared with 3 months in both groups. A possible explanation is that as the facial contour normalizes (losing the “moon face” as a result of weaning), the relationship between the mass and the rest of the facial structures changes, making hemangiomas more noticeable. Functionally, 6 of 8 patients had an improved visual function at 1 year despite lower VAS scores in patients with periorbital lesions compared with other facial areas. This suggests either that periorbital

**FIGURE 3**
Confidence interval (CI) plot of weight and height changes over time.
lesions are less steroid responsive or that their functional impairment is because of a deeper component that is not easily discernible on photographs.

We were also able to document biological changes in the circulating levels of angiogenic markers as result of intervention. Several factors, such as VEGF, VCAM-1, proliferating cell nuclear antigen, and interleukin 16 were implicated in the early proliferation phase.7,27–29 As hemangiomas involute, other growth factors seem to play a more significant role, such as bFGF.8 Two studies documented decreasing levels of proangiogenic factors, such as VEGF, with steroids and interferon.9,10 In our study, the only 2 markers that decreased significantly as the hemangiomas stopped proliferating were VCAM-1 and endoglin, suggesting that they are the most sensitive markers of endothelial proliferation and may be used as surrogate markers. The magnitude of the changes was more significant in the oral group than in the IV group correlating with the clinical response. The lack of changes noted in the VEGF values is possibly explained by the short duration of follow-up data (3 months). The serum and urine bFGF values did not change over the 3 months of follow-up as expected. Each infant served as his/her own control, considering that disease-specific and age-matched levels are ethically unjustified.

The safety data were similar to published studies.10,31 Parental reports of adverse events attributed to the medication were not different between the 2 study groups. The frequency of hypertension was 18.6% in the oral group and 13.1% in the IV group. Abnormal cortisol values were more frequently found in the oral group (78% vs 60%), with 32% of the abnormal tests in the oral group suggestive of severe adrenal suppression. This finding did not clinically translate into increased infections. The growth retardation was more pronounced in the oral group, similar to previous studies. Although our data suggest that the oral group had overall more adverse effects than the IV group, the sample size of our study was not sufficient to clearly establish a difference.

**CONCLUSIONS**

This randomized, blinded study showed that corticosteroids are efficacious in reducing the size of hemangiomas. Oral corticosteroids offered more clinical and biological benefit than the pulse steroids, however, with an increased risk of adverse effects. The pulse steroids may play a role in hemangiomas where a fast therapeutic response is needed (eg, eyelid hemangioma with complete visual axis occlusion) followed up by a shorter oral steroid regimen. Our data suggest that angiogenic proteins, such as VCAM-1 and endoglin, may have clinical use in monitoring clinical response and making therapeutic decisions. More prospective studies on the clinico-biological dose response of various corticosteroid regimens are needed.

**ACKNOWLEDGMENTS**

This study was awarded the 2002 William Weston Grant from the Society for Pediatric Dermatology.

We thank Dr Robert C. Pashby, pediatric ophthalmologist, who assessed all of the patients with visual compromise; dermatology nurses Lesley Eisel and Alejandra Stuparich and research assistants Susan Britton and Nicole Brown for navigating us smoothly through the study process; and Marg Mather for administrative support.

**REFERENCES**

ABSTRACT

OBJECTIVES. The purpose of this work was to compare the efficacy of propofol, a hypnotic agent, to the regimen of morphine, atropine, and suxamethonium as an induction agent for nonemergency neonatal endotracheal intubation. We hypothesized that propofol aids intubation by allowing the continuation of spontaneous breathing.

PATIENTS AND METHODS. We conducted a randomized, open-label, controlled trial of infants who required nonemergency endotracheal intubation. Primary outcome was successful intubation confirmed by chest auscultation and clinical examination of the infant.

RESULTS. Infants randomly assigned to propofol (n = 33) and the morphine, atropine, and suxamethonium regimen (n = 30) were comparable in median gestational age (27 vs 28 weeks), birth weight (1020 vs 1095 g), weight at intubation (1068 vs 1275 g), and age at intubation (4 vs 3 days). Sleep or muscle relaxation were achieved within 60 seconds in both groups, but time to achieve successful intubation was more than twice as fast with propofol (120 vs 260 seconds). Blood pressure and heart rates were not different, but intraprocedural oxygen saturations were significantly lower in infants on the morphine, atropine, and suxamethonium regimen (trough arterial oxygen saturation: 60% vs 80%). Nasal/oral trauma was less common, and recovery time was shorter (780 vs 1425 seconds) in the propofol group. No significant adverse effects were seen in either group.

CONCLUSIONS. Propofol is more effective than the morphine, atropine, and suxamethonium regimen as an induction agent to facilitate neonatal nasal endotracheal intubation. Importantly, hypoxemia was less severe, probably because of the maintenance of spontaneous breathing. A controlled environment may have
promoted the ease of intubation, resulting in less trauma. The shorter duration of action would be advantageous in a compromised infant.

Endotracheal intubation is a common but necessary procedure in the NICU. Newborn infants are unlikely to be different from older patients, where conscious intubation, regardless of the skill and expertise of the operator, is most likely to cause considerable discomfort and result in adverse physiologic sequelae, such as increased systemic and intracranial pressures, hypoxemia, and profound bradycardia. Studies have unequivocally demonstrated that premedicating infants who require intubation with various forms of induction agents increases the speed of successful intubation and reduces the likelihood of the occurrence of associated adverse sequelae.

However, there is no consensus as to which types of drugs are best for preintubation medications during the neonatal period. For example, a recent survey found that different combinations of premedication regimes were used in NICUs around the United Kingdom and Australia. Common combinations use an analgesic for pain, a vagolytic to counteract the reflex bradycardia associated with laryngeal stimulation, and a muscle relaxant to ameliorate struggling and to facilitate the intubation process.

A combination morphine, atropine, and suxamethonium regimen (MASux) is often used as a premedication for intubation. In our previous study, we found that MASux was superior to conscious intubation by increasing the speed and success rate of intubation and by decreasing the incidence of oropharyngeal trauma, a common complication of intubating an awake infant. MASux, however, is time consuming to prepare, especially when needed urgently, because many countries require dual accountability (eg, by 2 registered nurses) for the dispensation of narcotic agents. Furthermore, suxamethonium, a short-acting nondepolarizing muscle relaxant, may cause potentially serious adverse effects, such as hyperkalemia, profound vagotonia, malignant hyperthermia, masseter spasm, and systemic and ocular hypertension. Furthermore, complete and unexpectedly prolonged muscle relaxation and apnea secondary to paralytic agents, such as suxamethonium or pancuronium, may be detrimental if an airway is unable to be promptly secured.

Therefore, in this study, we sought to compare the efficacy of a single hypnotic agent, propofol (Diprivan), to MASux as an induction agent for nonemergency neonatal endotracheal intubation. Propofol is a soy-based formula commonly used in short operative procedures for adults and older children. Propofol, however, has not been explored for premedication purposes in newborn infants. We chose to explore propofol for this indication because of the convenience of its use as a single agent and its ability to preserve spontaneous respirations while providing hypnosis. We, therefore, hypothesize that propofol is more effective than MASux for nonemergency endotracheal intubation.

Methods
This was a randomized, controlled, open-label study conducted at the Department of Newborn Care at the Royal Hospital for Women in Randwick between March 2004 and December 2005. All of the newborn infants requiring elective or semielective (nonemergency) intubations were eligible if there was sufficient time to obtain informed parental consent. Infants admitted intubated at birth or by retrieval team would not be included unless there was a subsequent need for semielective intubation. Infants with major congenital abnormalities, whose parents had insufficient English-language skills to comprehend all of the explanations or who required emergency intubation (eg, resuscitation in the delivery suite), were excluded from the study. Infants requiring multiple intubations during their hospital stay were allowed to be in the study for only a single intubation. The institutional human research ethics committee approved the study before its commencement. The protocol was registered with the Interdisciplinary Maternal Perinatal Australasian Clinical Trials Network.

The primary purpose of this study was to compare the times required to achieve successful intubation, as well as to compare intraprocedural oxygen saturation, heart rate, and blood pressure changes between the propofol and MASux groups. Additional analyses were also performed on other related factors, as illustrated in Table 1. Random sampling numbers, based on a random number table, were used to assign each infant to blocks of 10 to receive either propofol or MASux after stratification by body weight (<1250 g and >1250 g) at the time of intubation. Group assignments were drawn from consecutively numbered, sealed, opaque envelopes that were opened by the trial team on the infant’s admission into the study immediately before intubation. Random sequences and envelopes were prepared by a senior nurse who was entirely uninvolved in the trial.

Blinding was not possible, because the drugs were very different in appearance: propofol is opaque and white, whereas MASux is a combination of 3 different ampoules of clear liquid. Morphine, an accountable restricted narcotic, was required to be prescribed for each individual patient, as well as to be prepared and accounted for by 2 registered nurses. This administration procedure was the main reason for a lengthy period of drug preparation in this regime. More importantly, the modes of action and the effects of these drugs were also quite different (muscle paralysis versus hypnosis) in assessing infants’ readiness for intubation.
Drug Doses and Administration

Propofol
The recommended lowest bolus induction dose for propofol in adults and children >3 years old is 2.5 mg/kg intravenously, with an infusion rate of 1.5 to 3.0 mg/kg per hour for the maintenance of anesthesia during an operative procedure. Safety and efficacy in newborn infants have not been well established, and, to date, infusions are not recommended for this group of patients (as development of lactic acidosis documented previously in older patients). We, therefore, chose to administer propofol as a single 2.5 mg/kg intravenous dose based on the above recommendations.

After a single intravenous dose of 2.5 mg/kg of propofol, eyelash reflex was tested by the intubator approximately every 10 seconds, the loss of which determined the onset of sleep or hypnosis. A maximum of 2 doses of propofol (2.5 mg/kg each) was allowed. The infant would then default to MASux if sleep had not been achieved in the desired time frame of 3 minutes or after the second dose of propofol.

MASux
This is the standard induction regime used in our NICU. The doses of drugs are as follows (all intravenous): morphine, 100 μg/kg; atropine, 10 μg/kg; and suxamethonium, 2 mg/kg. Two repeat doses of suxamethonium at 1 mg/kg each (maximum total dose of 4 mg/kg per intubation attempt) were administered if muscle relaxation was not achieved in the space of 3 to 5 minutes. Repeat applications of suxamethonium up to a maximum total of 4 mg/kg were allowed (if required) for each intubation attempt.

The Intubation Procedure and Personnel
For nonemergency intubation for prolonged ventilation, nasal intubation was the preferred route in our unit. Each infant was preoxygenated by positive-pressure mask ventilation (delivered by positive end-expiratory pressure-controlled Neopuff) and 100% oxygen before the administration of the induction agents. We do not have nurse practitioners in our NICU, and all of the intubations were conducted by medical officers of varying seniorities. As per our unit policy, each doctor was allowed a maximum of 2 intubation attempts, and each attempt was curtailed if the heart rate decreased below 60 beats per minute and/or the oxygen saturations decreased below 60%. The infant, after cessation of each attempt, was then reventilated by positive-pressure breaths with a mask delivering 100% oxygen. Intubation was then recommenced after reestablishment of the heart rate to >120 beats per minute and oxygen saturations to >90%.

The Royal Hospital for Women is a major teaching hospital for pediatric and neonatal subspecialty training. The majority of medical officers in our NICU are therefore seconded on 3-monthly periods from the adjacent Sydney Children’s Hospital. The first medical officer to attempt an intubation was usually a registrar (with ≥2 previous successful intubation attempts) or, less frequently, a resident who may be having the first hands-on intubation attempt, although with previous exposure to the procedure. Each intubation was personally supervised by doctors who were already technically competent in neonatal intubation (either a neonatal fellow or a consultant neonatologist), and the supervisor would then take over the intubation process after 2

<table>
<thead>
<tr>
<th>Variable</th>
<th>MASux (n = 30)</th>
<th>Propofol (n = 33)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gestation, median (IQR), wk</td>
<td>28 (25–31)</td>
<td>27 (25–30)</td>
<td>.393</td>
</tr>
<tr>
<td>Birth weight, median (IQR), g</td>
<td>1095 (759–1612)</td>
<td>1020.0 (770–1455)</td>
<td>.869</td>
</tr>
<tr>
<td>CRIB II score, median (IQR)</td>
<td>8 (3–11)</td>
<td>9 (5–12)</td>
<td>.372</td>
</tr>
<tr>
<td>Age at intubation, median (IQR), d</td>
<td>3.0 (1.0–33.8)</td>
<td>4.0 (1.0–16.0)</td>
<td>.778</td>
</tr>
<tr>
<td>Weight at intubation, median (IQR), g</td>
<td>1275 (905–1972)</td>
<td>1068 (810–1495)</td>
<td>.215</td>
</tr>
<tr>
<td>Male gender, n (%)</td>
<td>11 (36.7)</td>
<td>18 (54.5)</td>
<td>.251</td>
</tr>
<tr>
<td>Nasal ETT, n (%)</td>
<td>28 (93.3)</td>
<td>33 (100)</td>
<td>.431</td>
</tr>
<tr>
<td>First attempt intubator, n (%)</td>
<td>2 (6.7)</td>
<td>5 (15.2)</td>
<td>.504</td>
</tr>
<tr>
<td>Registrar</td>
<td>27 (90)</td>
<td>26 (78.8)</td>
<td>.387</td>
</tr>
<tr>
<td>Fellow</td>
<td>1 (3.3)</td>
<td>2 (6.1)</td>
<td>.945</td>
</tr>
<tr>
<td>Indications for intubation, n (%)</td>
<td>16 (53.3)</td>
<td>23 (69.7)</td>
<td>.284</td>
</tr>
<tr>
<td>RDS</td>
<td>0 (0)</td>
<td>1 (3.0)</td>
<td>.955</td>
</tr>
<tr>
<td>CLD</td>
<td>3 (10)</td>
<td>2 (6.1)</td>
<td>.912</td>
</tr>
<tr>
<td>NEC</td>
<td>4 (13.3)</td>
<td>2 (6.1)</td>
<td>.589</td>
</tr>
<tr>
<td>Preoperative</td>
<td>5 (16.7)</td>
<td>0 (0.0)</td>
<td>.048</td>
</tr>
<tr>
<td>Others*</td>
<td>2 (6.7)</td>
<td>5 (15.2)</td>
<td>.504</td>
</tr>
</tbody>
</table>

CLD indicates chronic lung disease; CRIB, Clinical Risk Index for Babies score II; ETT, endotracheal tube; NEC, necrotizing enterocolitis; RDS, respiratory distress syndrome; IQR, interquartile range.

* Includes meconium aspiration syndrome and hypoxic ischemic encephalopathy.
failed attempts from the junior officer. Nasal intubation is the preferred route in the unit. Oral intubations were performed only when emergency airway access was required, such as during resuscitation in the delivery suite or after failures of nasal intubations.

An intubation was considered successful if there was appreciable and bilaterally equal chest movement and air entry in conjunction with rising and stable oxygen saturations and heart rate after insertion of the endotracheal tube. After stabilization, each intubation was then confirmed by a portable chest radiograph, as per routine clinical management.

Timing of Procedures and Data Recording
A clinically uninvolved member of the team, also not a trial investigator, was designated as an observer, data recorder, and timekeeper. The following time periods were recorded: (1) medication preparation time: time from randomization until the medications were ready for administration; (2) time to achieve sleep and muscle relaxation: from administration of medications until onset of sleep/hypnosis (defined as loss of eyelash reflex) or muscle relaxation (defined as loss of voluntary muscle activity); (3) intubation time: from the first insertion of laryngoscope to clinical confirmation of successful intubation after all of the attempts were completed (clinical confirmation of successful intubation is evidenced by adequate chest movement, rising oxygen saturations, and bilaterally equal air entry on auscultation); and (4) recovery time: from onset of sleep and muscle relaxation (above) to return of spontaneous muscle movement.

Information recorded during the intubation process included the number of intubation attempts, additional doses of induction agents required, and the presence of intubation trauma, defined as the presence of blood in the nasal or oropharyngeal areas during or after intubation. Other data collected included head ultrasound appearances, lactate and base deficit levels before and after the intubation, and the severity of illness on admission (according to the Clinical Risk Index for Babies II score).14

The vital signs (including heart rate, oxygen saturations, and blood pressure) of the infants were continuously monitored on Spacelabs modules. Blood pressure was measured by either an indwelling intra-arterial cannula (usually umbilical arteries or radial arteries) or by an appropriately sized Dinamap blood pressure cuff that was inflated every 30 seconds before, during, and after the procedure. Oxygen saturations were also recorded simultaneously with a portable pulse oximeter (Radical, Masimo, Irvine, CA). Records of vital signs were commenced 5 minutes before administration of medications and continued for 5 minutes after complete recovery. Other than time intervals and baseline vital signs, the trough (lowest if any) heart rate, oxygen saturation, and blood pressure during the intubation process and the recovery period were recorded for comparisons.

A peripheral 24-G intravenous cannulae was established in each patient and flushed (to ensure patency) with 0.5 mL of normal saline before administration of medications. For each data set, the lactate and base deficit levels were categorized as either normal or increased ($\geq 2.2$ mmol/L and $\geq 2.0$, respectively), based on our laboratory reference values.

Sample Size and Statistical Analysis
The number of subjects required for this trial was calculated based on findings from our previous study.6 By selecting a power of 0.8 and a 2-tailed $\alpha$ of .05, 28 infants in each arm were necessary to demonstrate a 30% difference in the time to achieve intubation. Statistical analysis was undertaken on an intention-to-treat basis according to a preestablished analysis plan. This was to allow for possible crossovers to MASux from propofol, should the latter fail to achieve hypnosis.

Results are presented as percentages and medians with interquartile ranges (25th to 75th percentile). The Fisher’s exact and Mann-Whitney $U$ tests were used where appropriate. All of the computations were performed using SPSS 10 (SPSS Inc, Chicago, IL) and MedCalc for Windows 5.00.017 (MedCalc Software, Mariakerke, Belgium).

RESULTS
Sixty-seven infants needing semielective intubation were assessed for eligibility by the investigator team. Parents of 4 infants declined consents. Sixty-three infants (33 propofol and 30 MASux) were enrolled. The 2 groups were comparable in gestational ages, Clinical Risk Index for Babies II scores, birth weights, and weights and ages at intubation (Table 1). Most of the infants were intubated for respiratory distress syndrome.

Time Intervals
MASux took 5 times longer (median, interquartile ranges) to prepare than propofol (960, 900–1200 seconds vs 180, 180–210 seconds, respectively; $P < .001$). Figure 1 shows the time to achieve sleep and muscle relaxation, successful intubation, and recovery. The times to achieve sleep and muscle relaxation (60, 60–120 seconds vs 60, 30|60 seconds, respectively) for MASux versus propofol were comparable ($P = .087$). Successful intubation was more than twice as fast in the propofol group compared with MASux (120, 60|180 seconds vs 260, 60|435 seconds, respectively; $P = .007$). Stratifying the time of intubation by the number of intubation attempts showed that intubation times were similar for those who were intubated successfully on first attempt (60, 60–120 seconds vs 60, 52–120 seconds for the 20 propofol and 13 MASux successful first attempts, respectively; $P = .641$). Intubations that required $>1$
attempt were significantly faster in the propofol group than the intubation time required for the MASux (180, 120–300 seconds vs 360, 300–750 seconds, respectively; \(P < .001\)). Propofol infants regained spontaneous voluntary movements (ie, recovery) almost twice as fast as MASux (780, 360–1110 vs 1425, 645–2250 seconds, respectively; \(P = .002\)).

Details Regarding the Intubation Process
The majority of initial intubation attempts (84%) were performed by registrars who had previously had ≥2 successful intubation attempts (Table 1). Slightly more MASux infants (17 [57%] vs propofol: 13 [39%]; \(P = .263\)) required multiple attempts to achieve successful intubation (median, interquartile range: 2, 1–3 vs 1, 1–2, respectively; \(P = .082\)) and, thus, had to be deferred to senior doctors to complete the intubation process (13 [43%] vs 10 [30%]; \(P = .421\)).

Vital Signs
Heart rates (Fig 2) and oxygen saturations (Fig 3) decreased in both groups during the intubation process, whereas mean arterial pressures (Fig 4) increased in both groups. Differences in heart rates and mean arterial pressures were not significantly different, but oxygen saturations were significantly lower during intubation in the MASux group (Fig 3). Median (interquartile range) intraprocedural oxygen saturations were 60% (43%–82%) in the MASux group and 80% (67%–88%) in the propofol group (\(P = .019\)). Median oxygen saturations at recovery were also significantly higher in the propofol group (median, interquartile: 95%, 92%–98% vs 92%, 90%–96%; \(P = .008\)).

Adverse Events Related to the Induction Agents
Seven infants on MASux (23%) and 2 infants on propofol (6%) sustained intubation-related trauma (\(P = .117\)).
Serum lactate was analyzed before and after intubation as part of arterial blood gas measurements in 15 infants on MASux and 18 infants on propofol. Only 1 infant on MASux had an increase in lactate levels (>2.2 mmol/L). None of the infants became apneic after propofol induction, but 1 infant developed masseter spasm after suxamethonium administration. This complication was self-limiting, and intubation proceeded uneventfully after a period of manual insufflation. No other adverse effects were seen. No infant developed severe (grades 3 or 4) intraventricular hemorrhage after any of the intubation attempts.

**DISCUSSION**

Our study has shown that propofol, a hypnotic agent, is overall superior to MASux for the facilitation of elective neonatal endotracheal intubation. Propofol was also found to be more convenient and significantly faster to prepare than the combination of MASux, because it does not need dilution and is a single agent that is injected directly into the patient from a warmed ampoule. The speed of preparation of any induction agent is particularly important during emergencies. More importantly, intraprocudural oxygenation was maintained better in the propofol group, because spontaneous infant breathing was maintained. These favorable conditions facilitated the ease of the procedure so that intubation was achieved faster with less trauma.

Suxamethonium, a nondepolarizing neuromuscular blocker, is a commonly used paralytic for neonatal intubation induction because of its speed and short duration of action. It may have potential significant adverse effects, including vagotonia (and therefore, bradycardia), masseter spasm, rarely acute hyperkalemia, and malignant hyperthermia in genetically susceptible subjects, as well as increased systemic and intracocular pressures, prolonged muscle paralysis, and postadministration tetany. The dose of suxamethonium recommended for neonatal rapid sequence intubation is 2 to 3 mg/kg, a higher dose than that commonly used for
adults probably because the neonate has a dispro-
portionately larger volume of distribution and neuromus-
cular junction immaturity.\textsuperscript{17,18} Conversely, high doses of
suxamethonium may result in muscle rigidity, and lower
doses, such as 0.6 mg/kg, have been suggested to cir-
cumvent this problem.\textsuperscript{17} However, a number of infants
(7 of 30 infants [23\%]) required repeated doses to
achieve muscle relaxation in our study, and lower doses
of suxamethonium may not be as effective in the neo-
natal population. Masseter spasm, as witnessed in 1 pa-
tient in the MASux group, was not an issue in the
propofol group.

Regardless of the type of induction agent used, the
maximum benefits and attenuation of the adverse phys-
ologic responses of intubation would only be achieved if
the infant is adequately muscle relaxed before the pro-
cedure.\textsuperscript{18} Some agents that have been explored for this
purpose include thiopental,\textsuperscript{19} suxamethonium,\textsuperscript{6,20} pan-
curonium,\textsuperscript{2} and mivacurium,\textsuperscript{20} all of which have proved
to be superior to awake intubation. However, each of
these agents, as a consequence of their mode of action,
causes apnea, which may then result in profound hy-
poxemia if the operator is unable to promptly secure an
airway in the infant.

Nevertheless, many NICUs continue to practice
awake or conscious intubation for a variety of reasons.
First, there is no firm consensus as to which combination
of premedication is best, and many NICUs are located in
teaching hospitals where inexperienced operators may
not be able to promptly procure a patent airway. Anal-
gesia or sedatives such as morphine or benzodiazepines\textsuperscript{i5}
have been used alone to facilitate intubation without
muscle relaxation. However, even morphine has been
shown to be relatively ineffective in ameliorating the
pain response generated by supposedly innocuous pro-
cedures, such as heel pricks, in preterm infants.\textsuperscript{12} A
struggling infant may then paradoxically make the in-
tubation procedure more difficult for the inexperienced
operator. As illustrated in our setting, a teaching hospi-
tal, where the majority of intubators are relatively inex-
perienced, optimal intubation conditions may increase
the success rate of first-time intubations, and, indeed,
there was less need for assistance from senior doctors in
the propofol group.

We found that propofol induced satisfactory hypnosis
in less than a minute and that spontaneous respiration
was maintained throughout the intubation procedure at
the chosen intravenous dose of 2.5 mg/kg. The contin-
uation of spontaneous respiration may have allowed the
infants in the propofol group to maintain inappro-
dural oxygenation better than the suxamethonium
group. There was no precedent for this particular dosage
in the newborn (and especially low birth weight) infant,
and this dose was chosen because it was the lowest limit
of the range recommended for adult and pediatric pa-
tients.\textsuperscript{3} The treatment failures with this dose was com-
parable to the MASux group. We also found that a single
dose of propofol was adequate to complete the majority
of intubation attempts without adverse effects that have
been noted in adults, such as hypotension, vagotonia, or
apnea.\textsuperscript{9} In addition, none of the infants suffered intra-
ventricular hemorrhage. Lactic acidosis, an adverse ef-
effect of prolonged infusions in older patients, was not
noted, but we would strongly caution against using
propofol as an infusion in newborn infants until further
evidence is available, because severe complications, such
as zinc deficiency, metabolic acidosis, rhabdomyolysis,
hyperkalemia, renal failure, and death have been re-
ported in adult patients administered with propofol in-
fusions >5 mg/kg per hour for >58 hours.\textsuperscript{21}

Unfortunately, neither of these regimes may be ad-
ministered without intravenous access. Propofol, in par-
ticular, may cause severe pain if it extravasates into
tissues.\textsuperscript{9} There is, therefore, a need to explore agents that
may be administered by nonintravenous routes in case
patent venous access cannot be obtained. For example,
ketamine can be administered intramuscularly,\textsuperscript{9} and ni-
trous oxide can be inhaled\textsuperscript{22} (although the latter requires
special dispensing equipment).

Therefore, until evidence for nonintravenous induc-
tion agents is available, we strongly condone the use of
intravenous premedications for this purpose, because
the noxious effects of conscious intubation are well doc-
umented and undisputed. Although adults and older
children who require this procedure are often anesthe-
tized or sedated, similar practice is not routine in the
newborn population, and, indeed, in a 10-day prospec-
tive survey of 75 neonatal and PICUs in France, Simon et
al\textsuperscript{13} found that only 37.1\% of neonates (as opposed to
67.3\% of infants and 91.7\% of children) received pre-
medication before intubation, and premedication was
particularly infrequently used for the youngest and
smallest infants, who are theoretically at greatest risk of
developing adverse physiologic sequelae, such as intra-
ventricular hemorrhage.

This may be because of a misconception that newborn
infants are less susceptible to pain when compared with
older infants or adults, but substantial evidence shows
that even fetuses respond adversely to noxious stimuli
by 26 weeks’ gestational age.\textsuperscript{24,25} that thalamocortical
connections responsible for pain perception are present
from 26 weeks’ gestation,\textsuperscript{24} and that repeated noxious
stimuli may cause longer-term behavioral changes.\textsuperscript{26}
Another reason may be because the operator, with ad-
equate assistance, can easily overcome the struggling
efforts of small newborn infants during a noxious pro-
cedure. A further reason may be, as discussed previ-
ously, a natural trepidation to use muscle relaxation in
case a patent airway is unable to be obtained.

Our study is, therefore, the first to compare a hyp-
notic to a combination of analgesics and muscle relax-
ants for newborn semielective endotracheal intubation,
and we have found that propofol is superior to the MASux as an induction agent to facilitate the procedure. Propofol also offered additional advantages, such as maintenance of spontaneous respiration, less profound hypoxemia, and less procedure-related trauma compared with MASux. Faster recovery could also be an advantage in a compromised infant or in a case of difficult intubation. We, therefore, recommend further large-scale studies with adequate sample size to assess the short- and long-term safety of propofol, especially for more urgent neonatal endotracheal intubations.

REFERENCES

A Randomized, Controlled Trial of a Removable Brace Versus Casting in Children With Low-Risk Ankle Fractures

Kathy Boutis, MDa,b, Andrew R. Willan, PhDb, Paul Babyn, MDb, Unni G. Narayanan, MDb,d, Benjamin Alman, MDb, Suzanne Schuh, MDa,b

aDivision of Emergency Medicine, bPopulation Health Sciences, Research Institute, and Divisions of cDiagnostic Imaging and dOrthopedic Surgery, Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada

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

ABSTRACT

OBJECTIVES. Isolated distal fibular ankle fractures in children are very common and at very low risk for future complications. Nevertheless, standard therapy for these fractures still consists of casting, a practice that carries risks, inconveniences, and use of subspecialty health care resources. Therefore, the main objective of this study was to determine whether children who have these low-risk ankle fractures that are treated with a removable ankle brace have at least as effective a recovery of physical function as those that are treated with a cast.

METHODS. This was a noninferiority, randomized, single-blind trial in which children who were 5 to 18 years of age and treated in a pediatric emergency department for low-risk ankle fractures were randomly assigned to a removable ankle brace or a below-knee walking cast. The primary outcome at 4 weeks was physical function, measured by using the modified Activities Scale for Kids. Additional outcomes included patient preferences and costs.

RESULTS. The mean activity score at 4 weeks was 91.3% in the brace group (n = 54), and this was significantly higher than the mean of 85.3% in the cast group (n = 50). Significantly more children who were treated with a brace had returned to baseline activities by 4 weeks compared with those who were casted (80.8% vs 59.5%). Fifty-four percent of the casted children would have preferred the brace, but only 5.7% of children who received the brace would have preferred the cast. The cost-effectiveness acceptability curve was always >80%; therefore, the brace was cost-effective compared with the cast.

CONCLUSIONS. The removable ankle brace is more effective than the cast with respect to recovery of physical function, is associated with a faster return to baseline activities, is superior with respect to patient preferences, and is also cost-effective.
ANKLE INJURIES ARE very common among children. However, most ankle injuries are minor with little need for imaging or for subspecialty care. They include sprains and fractures such as isolated distal fibular fractures, which are the most common fractures of the lower extremity. These fractures are stable and at negligible risk for premature closure of the growth plate because fractures through fibular growth plate typically spare the zone of proliferation where the growth occurs. Despite their benign natural history, these stable, low-risk fractures are often immobilized in a cast for 3 to 4 weeks. The main disadvantages of casting include unpleasant and prolonged immobilization and the inconvenience of return visits to an orthopedic facility. Moreover, casting may be unnecessary and expensive and may have complications that potentially are greater than those that are associated with the fracture itself. Therefore, a comparison of this traditional treatment with a more convenient method is imperative. The single study that found that children who had Salter-Harris I fractures of the distal fibula and were treated with a tensor bandage returned to normal activities sooner than did those who were treated in a cast was limited by the use of unvalidated outcome measures that were assessed by investigators who were not blinded to the treatment allocation. Therefore, solid evidence to support the use of a more cost-effective alternative form of immobilization that offers comparable comfort and return to activities is lacking.

Our study challenged the current practice of routine casting and the necessity for orthopedic care. This is the first randomized, controlled trial to compare a removable ankle brace with conventional rigid cast immobilization for low-risk ankle fractures in children in terms of recovery of physical function. In addition, because the brace may result in lower morbidity and reduced use of health care resources compared with casting, a formal cost-effectiveness analysis was a part of this trial.

FIGURE 1
Radiographs of distal fibular fracture types included in the study.
Recruitment and Baseline Assessment
Before the study, pediatric emergency medicine (PEM) staff physicians and research assistants were instructed by the participating orthopedic surgeons in the physical examination of the ankle and the accurate diagnosis of the eligible ankle injuries. The triage nurse alerted the research team about all ankle injuries that presented to the ED between the hours of 8:00 and 12:00 AM. These children were assessed by both the research nurse and staff PEM physicians. The patients were sent for the standard ankle (anteroposterior, lateral, and mortise) radiographs. The on-call radiologist analyzed the radiographs to record whether the growth plate was open or closed and for the presence of soft tissue swelling, effusions, and fractures. The final diagnosis was made by the PEM staff physician on the basis of clinical examination and the radiologist’s report. When a child met eligibility criteria, informed consent and assent (children ≥ 7 years) was obtained. A standardized study sheet that contained the clinical data and a detailed diagram of the ankle (prepared a priori by the investigative team) was completed for each child. The time and mechanism of injury, ability to bear weight, sociodemographic factors, and physical examination findings were recorded. Physical examination findings included the site of maximal pain and tenderness, the presence and the location of swelling, and degree of pain with weight bearing. Patients also completed the modified performance Activities Scale for Kids (ASKp) (refs 13-16; N. L. Young, PhD, written communication, 2002) to assess baseline physical function in the week before the injury.

Randomization
Recruited patients were randomly assigned to receive either an ankle brace or a below-knee fiberglass cast. Concealed treatment allocation was provided by an online randomization program (www.randomize.net) using block randomization with random block sizes of 6 and 8, with immediate e-mail notification of treatment group to the research coordinator.

Study Interventions
A trained research cast technician placed either a below-knee fiberglass walking cast or the Aircast (Vista, CA) Air-Stirrup ankle brace in accordance with the patient’s group assignment. Patients were instructed to wear a sock and a shoe in conjunction with the brace. All patients were provided with crutches. After a period of no weight bearing for 5 days in their respective immobilization devices, both populations were instructed to bear weight as tolerated and to transition to full, unassisted weight bearing and activities if there was no pain. The patients in the brace group were instructed to remove the Air-Stirrup when comfortable with ambulating after the initial 5-day period. Apart from specific instructions regarding care of the cast or the ankle brace, both populations received identical instructions on initial elevation and non-weight-bearing guidelines, type and frequency of analgesics, and reasons to return for medical attention before scheduled follow-up appointments. Study patients were not scheduled for any follow-up radiographs.

Patients received weekly follow-up telephone calls over 4 weeks to address parental concerns and remind families to complete the expense and clinical diaries provided in the ED. The clinical diary was used to record the actual amount and frequency of analgesia, weekly pain scores, and weekly return to baseline activities. The expense diary included any costs incurred by the family during a 4-week period related to the index ankle injury. Children who were treated with cast immobilization were provided an appointment to return 3 weeks after injury to our orthopedic clinic for removal of the cast.

Within 1 week of patient recruitment, a single staff radiologist who had expertise in pediatric musculoskeletal imaging and was blinded to the details of the initial presentation and initial radiology report read all study radiographs to ensure consistency of interpretation and to ascertain that there were no diagnostic errors at the ED visit. Differences of opinion with respect to the final radiographic diagnosis were resolved by consensus after review of the clinical history and radiographs with the collaborating orthopedic surgeon and the participating radiologist.

Four weeks after the injury, all patients were visited at home by the research physiotherapist, who was blinded to the treatment allocation. To preserve blinding, patients and families were instructed not to mention to the visiting physiotherapist which immobilization device the patient received. All patients were also provided with an opaque stocking that was placed on the affected leg before the physiotherapist’s assessment. The physiotherapist ensured completion of the 4-week assessments and collected the completed diaries. Three months after injury, the patients were telephoned by a blinded research assistant to assess subsequent complications.

Outcome Measures
The primary outcome measure was the modified ASKp score (refs 13-16; N. L. Young, PhD, written communication, 2002) at 4 weeks. The modified ASKp is an assessment of a child’s current physical function that is based on activities performed during the week before completion of the questionnaire. The questionnaire has 30 items that reflect clinician and child perspectives on pediatric daily activities.16 The ASKp has been found to be highly reliable,13,14 have excellent construct validity, and be responsive to change.13,14,16,17 The ASKp has been validated in 5- to 15-year-old children with fractures and other musculoskeletal problems.16 Because only 70% of the questions in the ASKp pertain directly or indirectly to ankle function, 8 additional questions re-
lated to ankle activity were included in consultation with the author of the ASKp to create a modified ASKp score. Secondary outcome measures also obtained at the 4-week follow-up visit included range of motion as measured by a certified physiotherapist using a goniometer, pain with walking using the validated Bieri Face Pain Scale-Revised, return to baseline activities as reported by parents, and patient preferences for 1 immobilization device versus the other.

Data on health care resources used included over-the-counter medications used, physician visits, ankle radiographs, casting (cast changes/replacements), and use of ankle braces. Total patient health care costs were estimated by multiplying health care resource use data by corresponding unit prices. Price sources included a survey of pharmacies, Ontario Schedule of Physician Benefits for physician fees and procedures, and manufacturers for cast materials and ankle brace. Additional costs included travel, child care expenses, and parental lost productivity. Cost for lost productivity was determined by multiplying $20.80, the average hourly wage of all employees between the ages of 24 and 54 from the 2005 Canadian Labor Force survey, by the number of hours reported off work. All costs and unit prices are given in Canadian dollars.

Statistical Analysis
All analyses were by intention to treat. The sample size was determined using methods that are appropriate for noninferiority trials (ie, trials with nonzero null hypotheses). The sample size of 111 patients was based on testing the null hypothesis ($H_0$) that the brace is $\leq 5\%$ less effective than a cast on the basis of the ASKp scores at 4 weeks, at the 5% level, and having an 80% probability of rejecting $H_0$ if brace and cast are equally effective. A 5% difference in effectiveness was chosen because it is approximately the difference in average ASKp scores between normal and mildly disabled patients (N. L. Young, PhD, written communication, 2002). In addition, it was assumed that $\sim 10\%$ of patients would be lost to follow-up. $H_0$, based on the week 4 ASKp scores, was tested by a $t$ test for a nonzero difference. For the other clinical outcomes, traditional 0 $H_0$s were tested. Proportions were compared with Fisher’s exact tests, means were compared with Student’s $t$ tests, and Cochrane test trend was used to compare treatment arms with respect to the ordered categorical outcome of patient satisfaction.

For the cost-effectiveness analysis, mean costs by treatment arm and between-treatment arm differences in mean costs were estimated assuming a $\gamma$ distribution to accommodate right skewing. Estimation was performed using Markov-chain Monte Carlo methods as facilitated by the software WinBUGS with vague priors (WinBUGS 1.4; MRC Biostatistics Unit, Cambridge, United Kingdom). To conduct a cost-effectiveness analysis, we performed the simultaneous estimation of the between-treatment arm differences in mean costs and effectiveness (week 4 ASKp scores), along with the corresponding variances and covariances, using Markov-chain Monte Carlo methods as facilitated by the WinBUGS software with vague priors.

RESULTS

Patient Recruitment and Baseline Characteristics
During the study period, 679 ankle injuries presented to the ED, 607 of which were screened. Of the 72 missed patients, 5 presented outside study recruitment times, and in the remaining cases, the ED staff did not alert study personnel. Of these, 12 would have been eligible for recruitment and 60 met exclusion criteria. Of the 607 screened patients, 475 were ineligible. Twenty-one of the 132 eligible patients declined participation. Of the 111 children who were enrolled and randomly assigned, 104 were included in the final analysis (Fig 2). Baseline characteristics of the study groups are summarized in Table 1.

Follow-up
Follow-up of the primary outcome, ASKp at 4 weeks, was completed in 103 of (99.0%) 104 of the children who were included in the final analysis. Ninety-nine (95.2%) of the patients completed the clinical diary, and 98 (94.2%) completed the expense diary. Ninety-four (90.4%) of the patients were reached for the 3-month follow-up telephone call.

Functional Outcomes
The mean ASKp ($\pm SE$) at 4 weeks was $91.3\%$ ($\pm 1.14\%$) in the brace group compared with $85.3\%$ ($\pm 2.06\%$) in the cast group. The difference between the mean ASKp (brace) and the mean ASKp (cast) was therefore 6% in favor of the brace, and the lower bound of the 95% confidence interval for this difference was 1.13%. As a result, the $H_0$ that the mean ASKp (brace) is inferior by 5% or more can be rejected ($P < .0001$). Furthermore, because the lower bound of the 95% confidence interval is positive, the hypothesis that the brace is inferior by any amount can also be rejected at the 5% level.

Although there were no differences in the 2 groups in pain, ability to bear weight, and range of motion at 4 weeks, a significantly greater proportion of children in the brace group compared with their casted counterparts had returned to baseline activities by this time (Table 2). At 3 months, there were no differences in pain, other residual symptoms such as swelling, and activity levels.

Immobilization Devices
Four of the 50 children in the cast group had their casts removed prematurely at 2 weeks. Two were placed in a brace, and 2 had no additional immobilization. These children were treated in the orthopedic fracture clinic by
residents who were unfamiliar with the study and who therefore chose management independent of study protocol. Figure 3 demonstrates brace usage during the 4-week period. There were no crossovers from the brace to the cast group.

Complications and Co-interventions
Twenty patients had unscheduled visits to a health care provider: 16 (32.0%) in the cast group and 4 (7.4%) in the brace group. Patients visited at least 1 of the following: primary care provider (5 casts, 2 braces), cast clinic (4 casts, no braces), ED (3 casts, 1 brace), physiotherapist (3 casts, no braces), and/or a sports medicine clinic (1 cast, 1 brace). Reasons in the cast group included poor

<table>
<thead>
<tr>
<th>Variable</th>
<th>Brace (n = 54)</th>
<th>Cast (n = 50)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD), y</td>
<td>10.55 (2.91)</td>
<td>9.94 (2.43)</td>
</tr>
<tr>
<td>5–10, n (%)</td>
<td>26 (48.10)</td>
<td>30 (60.00)</td>
</tr>
<tr>
<td>11–17, n (%)</td>
<td>28 (51.90)</td>
<td>20 (40.00)</td>
</tr>
<tr>
<td>Baseline modified ASKp (SE)</td>
<td>90.8 (1.36)</td>
<td>89.9 (1.70)</td>
</tr>
<tr>
<td>Baseline pain score (SE) out of 5</td>
<td>3.46 (0.16)</td>
<td>3.46 (0.15)</td>
</tr>
<tr>
<td>No weight bearing, n (%)</td>
<td>34 (62.90)</td>
<td>29 (58.00)</td>
</tr>
<tr>
<td>Partial with a lot of pain, n (%)</td>
<td>20 (37.10)</td>
<td>21 (42.00)</td>
</tr>
<tr>
<td>Diagnoses, n (%)</td>
<td>Rule out Salter-Harris I</td>
<td>36 (66.70) 39 (78.00)</td>
</tr>
<tr>
<td></td>
<td>Salter-Harris II</td>
<td>9 (16.70) 4 (8.00)</td>
</tr>
<tr>
<td></td>
<td>Avulsion of distal fibula</td>
<td>4 (7.50) 5 (10.00)</td>
</tr>
<tr>
<td></td>
<td>Avulsion of distal fibular epiphysis</td>
<td>5 (9.30) 2 (4.00)</td>
</tr>
</tbody>
</table>

| TABLE 2 Comparison of Secondary Outcomes at 4 Weeks |
|-----------------------------------------------|----------------|--------------|
| Parameter                                    | Brace (n = 54) | Cast (n = 50) |
| Pain with walking, n*                        | 54             | 50           |
| Mean FPS score (SE)                          | 0.32 (0.10)    | 0.33 (0.13)  |
| Able to bear weight fully with no pain, n*  | 54             | 52           |
| Return to baseline activities, n*            | 52             | 42           |
| Patient satisfaction with device, n*         | 53             | 50           |
| FPS indicates revised Face Pain Scale        |                |              |
| *Difference = −0.0126 (95% confidence interval [CI]: −0.028 to 0.363). |
| Relative risk (RR, defined as probability of outcome on brace/probability of outcome on cast) = 0.962 (95% CI: 0.776 to 1.19). |
| RR = 1.36 (95% CI: 1.03 to 1.80).            |                |              |
cast fit, itchy leg, and strength and range-of-motion issues. In the brace group, 1 patient developed a rash on the leg (likely as a result of not wearing socks with the brace), and the remaining 3 children had strength and range-of-motion issues. There were no serious adverse events in either group.

Patient and Parent Satisfaction and Preference
Patients and parents were asked at the 4-week visit how satisfied they were with the immobilization device and whether at any point they would have preferred an alternative device. Twenty-eight (52.8%) patients in the brace group rated themselves as “very happy” compared with 9 (18%) “very happy” patients in the cast group. Only 3 (5.7%) in the brace group reported that they would have preferred a cast, whereas 27 (54%) in the cast group reported that they would have preferred a brace. It is interesting that although there were no differences in parental preference, parental satisfaction with the brace, as measured by report of satisfaction on a 4-point Likert scale, was significantly greater than that reported in the cast group ($P < .0001$; Table 2).

Cost-effectiveness
The between-treatment cost comparisons are reported in Table 3. Total costs do not equal the sum of the 3 components because of the nonlinearity of the estimation procedure based on the $\gamma$ distribution. Total costs and health care costs were observed to be lower in the brace arm, whereas parental work loss costs and other costs were higher in the brace arm. Health care cost differences between the 2 interventions achieved statistical significance ($P < .0001$), favoring the brace. The cost-effectiveness acceptability curve is a plot of the probability that the brace is cost-effective as a function of the willingness to pay for a unit of effectiveness.30,31 Because the brace is observed to be more effective and less costly and the cost-effectiveness acceptability curve was always $>80\%$, the brace is observed to be cost-effective compared with cast.

**DISCUSSION**
This study is the first to demonstrate that a removable ankle brace is more effective than the below-knee walking cast with respect to recovery of physical function in children with low-risk ankle fractures. In addition, the brace was observed to be cost-effective and superior with respect to return to baseline activities and patient preferences. The brace can initially be applied readily by any health care provider and applied or taken off easily by older children or an adult family member and therefore does not require specialty orthopedic involvement, which is usually necessary for a cast. The convenience and efficiency of the brace may reduce costs, avoid ED visits for patients who receive a diagnosis of these injuries in a primary care setting, reduce time spent in the ED, and obviate the need for follow-up visits in busy general or pediatric orthopedic clinics.

The previously determined difference of the mean ASKp for “normal” versus “mild” disability is $\sim 5\%$ (N. L. Young, PhD, written communication, 2002). Therefore, the mean score difference of 6% in favor of the brace is clinically significant. In addition, the absolute mean score in the brace group was close to “normal” level,

---

**TABLE 3** Estimated Means and SEs for Costs According to Treatment Arm

<table>
<thead>
<tr>
<th>Costs</th>
<th>Brace Mean</th>
<th>SE</th>
<th>Cast Mean</th>
<th>SE</th>
<th>Brace – Cast Mean</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total costs</td>
<td>278.30</td>
<td>41.57</td>
<td>322.40</td>
<td>27.93</td>
<td>-44.09</td>
<td>50.10</td>
</tr>
<tr>
<td>Health care costs</td>
<td>90.88</td>
<td>4.15</td>
<td>156.60</td>
<td>10.87</td>
<td>-65.76</td>
<td>11.61</td>
</tr>
<tr>
<td>Parental work loss</td>
<td>149.60</td>
<td>47.19</td>
<td>121.10</td>
<td>17.11</td>
<td>28.52</td>
<td>50.24</td>
</tr>
<tr>
<td>Other</td>
<td>43.21</td>
<td>9.06</td>
<td>40.00</td>
<td>5.60</td>
<td>3.21</td>
<td>10.68</td>
</tr>
</tbody>
</table>

Values are Canadian dollars.
whereas that of the casted counterparts more closely correlated with "mild" disability (ref 33; N. L. Young, PhD, written communication, 2002). Therefore, our results achieved both statistical and clinical significance.

Our results are consistent with the findings by Gleeson et al.,11 which showed that children who had Salter-Harris I fractures of the distal fibula and were treated with a tensor bandage returned to baseline activities ~7 days sooner than those in the below-knee walking cast. Our study used a reliable and well-validated primary outcome measure10 that was measured by an outcomes assessor who was blinded to the treatment group. Furthermore, in addition to the Salter-Harris I fracture of the distal fibula, our study included Salter-Harris II and avulsion fibular fractures.

The absence of any serious adverse events in either group is consistent with the knowledge about isolated distal fibular fractures.3,5,34 The long-term complication of growth arrest in these fractures is very rare,3,6,7 and even if it occurs, there is no evidence to support that casting prevents this outcome or that it results in any functional impairment of the ankle.1 The complications that did occur were minor and could easily be assessed by a primary care provider with subsequent referrals to a physiotherapist or sports medicine clinic as required. Therefore, if the brace is used, then the routine referral of these low-risk ankle fractures to an orthopedic surgeon becomes unnecessary.

This study has 2 main limitations. First, our results may not be generalizable to physicians with limited expertise in diagnosing low-risk ankle fractures in children. Second, it is uncertain whether the high use of the brace in the early period of treatment reflected the increased vigilance that patients received during the study as opposed to what may happen in reality in the absence of weekly follow-up.

CONCLUSIONS
In children with low-risk ankle fractures, the removable brace is more effective than the cast with respect to an earlier recovery of physical function, is associated with greater patient/family satisfaction, and is less costly than the traditional casting.

ACKNOWLEDGMENTS
This study was supported by a grant from Physician Services Inc.

We acknowledge Dr Martin Pecaric of Contrail Consulting Services for data entry and graphic design of figures in the manuscript. Also, we thank Aircast for providing, free of charge, the Air-Stirrup ankle braces that were used in the study.

REFERENCES
Birth Weight, Postnatal Growth, and Risk for High Blood Pressure at 7 Years of Age: Results From the Collaborative Perinatal Project

Anusha H. Hemachandra, MD, MPH, Penelope P. Howards, PhD, Susan L. Furth, MD, PhD, Mark A. Klebanoff, MD, MPH

OBJECTIVE. A physiologic predisposition toward hypertension is theorized to result from the combination of intrauterine growth restriction followed by rapid catch-up growth. The objective of this study was to evaluate the effects of birth weight and weight gain during childhood on the risk for high blood pressure in childhood and to identify discrete periods of catch-up growth that put children with intrauterine growth restriction at increased risk for the development of high blood pressure later in life.

METHODS. The US Collaborative Perinatal Project (1959–1974) studied 55,908 pregnancies in an observational cohort at 12 medical centers in the United States and followed the offspring through 7 years of age. All white or black children who were born at term and completed the follow-up without kidney or heart disease were included in this posthoc analysis. z scores were calculated for weight at birth, 4 months, 1 year, 4 years, and 7 years on the basis of study means and SD. Changes in z scores were calculated for each interval.

RESULTS. Each 1-kg increase in birth weight increased the odds for high systolic blood pressure by 2.19 and high diastolic blood pressure by 1.82 when race and change in weight z scores were also included in the regression model. An increase in weight z score of 1 SD above the previous weight z score increased the odds for high systolic blood pressure at 7 years by 1.65 (birth to 4 months), 1.79 (4 months to 1 year), 1.71 (1–4 years), and 1.94 (4–7 years) in the full model. White race increased the odds for high systolic blood pressure by 1.51.

CONCLUSIONS. In this large biracial US cohort, infants who were small for gestational age were not at increased risk for high blood pressure at 7 years of age. However, children who crossed weight percentiles upward during early childhood did demonstrate an increased risk.
The “developmental origins of adult disease” hypothesis, also known as “fetal programming,” has been widely recognized as a possible mechanism for the development of a number of chronic diseases of adulthood. The hypothesis suggests that intrauterine compromise, leading to low birth weight, results in permanent alterations of fetal physiology that persist into the postnatal period. These adaptations confer a survival advantage on the fetus while in the suboptimal intrauterine milieu, but they are deleterious to the individual after birth, when nutrients and other resources are abundant. The hypothesized consequence is that these growth-restricted neonates grow into adults with an increased risk for chronic diseases such as cardiovascular disease, type 2 diabetes, metabolic syndrome, and osteoporosis.

Because fetal growth restriction (the failure of a fetus to achieve its own growth potential) can occur in infants of any weight, the risk for the development of chronic disease is not limited to the smallest infants. There is a continuum of risk across the birth weight spectrum. That risk seems to be compounded when intrauterine growth restriction (IUGR) is coupled with rapid postnatal catch-up growth. Several published reports have demonstrated the increased risk for hypertension, type 2 diabetes, metabolic syndrome, and obesity associated with the interaction between birth weight and postnatal growth.

Most of the articles that evaluated the impact of postnatal growth on subsequent health defined growth as the total amount of weight gained between birth and a second point in time. However, this crude measure of growth fails to take into account the increasing variation in weight of children as they grow and is not how pediatricians evaluate growth in normal children. For example, a girl who tracks along the 75th percentile of weight will weigh 10 kg at 1 year of age and 25 kg at 7 years of age, reflecting a weight gain of 15 kg. Another girl who tracks at the 25th percentile of weight will weigh 9 and 20 kg at 1 and 7 years of age, respectively, reflecting a weight gain of 11 kg. Although the first girl gained 4 kg more than the second girl, neither girl changed her relative weight compared with her peers. In this report, we propose to define “catch-up growth” not by absolute change in weight but rather by change in relative weight compared with other children and examine the association among birth weight, catch-up growth, and blood pressure (BP) at age 7 in a large cohort of American children.

METHODS
The Collaborative Perinatal Project (CPP) enrolled pregnant women at 12 academic medical centers in the United States in a nationwide cohort between 1959 and 1965. Women were enrolled at their first prenatal visit and were followed during pregnancy, labor, and delivery. The offspring were followed for 7 years, with multiple questionnaires regarding medical and social history and detailed neuropsychological testing. Vital signs including BP were recorded before neuropsychological testing at the age 7 visit, with a manual sphygmomanometer on the right arm of the child in a sitting position. Comprehensive descriptions of the method of the study have been published previously. Of note, the diastolic BP (DBP) measurement was taken by the fourth Korotkoff sound, as was standard procedure before 1977. The data are available for public use with patient identifiers omitted from the data set.

Of the 58,960 pregnancies enrolled in the study, 51,540 mothers of white or black race were identified. We excluded all mothers who were identified as Hispanic, Asian, or other ethnicity because they composed such a small proportion of the total population. After exclusion of stillbirths, terminations, preterm births, and women who dropped out of the study before delivery, 41,413 infants were born between 37 and 42 completed weeks of estimated gestational age by menstrual dating. Of these infants, 417 died before 7 years of age, leaving 40,996 eligible children. By the end of the study, 29,973 (73%) completed the 7-year follow-up and were eligible for inclusion in this analysis.

With the use of Tukey’s severe outlier criteria, as well as exclusion of data points that were 4 or more SDs from the mean, biologically implausible data were removed from the data set. These criteria were applied to birth weight, head circumference, chest circumference, birth length, placental weight, SBP and DBP at 7 years of age, and weight and height at 7 years of age. Children who had a diagnosis of heart or kidney disease (n = 109) were also excluded, resulting in a final study population of 29,710 children.

Small for gestational age (SGA) was defined as a birth weight <10th percentile for gender, race, and gestational age using birth weight distributions based on the 29,710 children, large for gestational age (LGA) was defined as birth weight >90th percentile, and all other infants were considered appropriate for gestational age (AGA). High BP was defined as SBP or DBP >90th percentile, as recommended by the 1996 Task Force Report on High Blood Pressure in Children and Adolescents. BP distributions for this study population, stratified by race and gender, were calculated, and this internal standard was used to identify children who were above the 90th percentile for SBP, DBP, or pulse pressure (PP). Maternal characteristics that are known to influence child BP were examined, including education and socioeconomic status, smoking, diabetes, and hypertension.

Childhood weight was recorded consistently at 5 times during the CPP follow-up: birth, 4 months, 1 year, 4 years, and 7 years. At each of these points, we calcu-
related $z$ scores for each recorded weight, based on study means and SDs:

The change in $z$ score was calculated for each individual, generating 4 interval changes in $z$ scores for each child. For example, a child who was at the 50th percentile at 4 months of age but crossed percentiles at 1 year of age would have a change in $z$ score of +1 for the interval of 4 months to 1 year. If that child continued to grow along the 84th percentile for the rest of the study, then the subsequent changes in $z$ score all would be 0, because the child’s relative position on a growth chart no longer changed.

When a weight measurement was missing for any individual, the change in $z$ score for that interval could not be calculated and the individual was not included in the regression analyses. These changes were used to assess body size in relation to the previously recorded size, thereby quantifying increase or decrease in relative size. These changes were included in multivariable logistic regression models, with birth weight and race, to predict high SBP, DBP, and PP at 7 years of age.

A preliminary logistic regression model with birth weight, race, change in weight $z$ score, and predictors of IUGR (smoking, poverty, and anemia) was also created, but smoking, poverty, and anemia were not statistically significantly associated with SBP, DBP, or PP, so they were dropped from the model. Forward stepwise logistic regression technique was used, with an entry criterion of $P < .05$ and a removal criterion of $P > .10$. The models were also run on the same population stratified by birth size. A total of 24 055 infants were AGA, 2802 infants were SGA, and 2853 infants were LGA. Another set of multivariable logistic regression models were run with interaction terms between birth weight and change in weight $z$ scores for each interval to determine whether size at birth and postnatal crossing of growth percentiles had a synergistic association with high BP at age 7. Statistical analysis was performed using SPSS version 11.0 software (SPSS, Chicago, IL).

**RESULTS**

Children who were born into this study population had a mean birth weight of $3.24 \pm 0.48$ kg and a gestational age of $39.7 \pm 1.4$ weeks on the basis of last menstrual period. Mean SBP was $102.1 \pm 10.2$ mm Hg, mean DBP was $61.3 \pm 9.8$ mm Hg, and mean PP was $40.8 \pm 10.3$ mm Hg.

Maternal characteristics with a potential influence on child BP are listed in Table 1. The women in this cohort were relatively young (mean: 24.5 ± 6.1 years) and thin (prepregnancy BMI 22.9 ± 4.3), with 46.7% reporting smoking and 5.4% reporting a diagnosis of hypertension at the time of presentation for prenatal care. The CPP collected data on “toxemia” rather than preeclampsia, and in this cohort, 2.8% were considered to have toxemia. During pregnancy, 14.2% of women were found to have a hematocrit <30%. Slightly more than half of these women lived below the federal poverty level, established by the US Census Bureau in 1960, and the mean number of years of education was 10.9 ± 2.4. Both black and white women were well represented in this group, with 47.7% identifying themselves as black and 52.3% identifying themselves as white.

Table 2 shows the Pearson bivariate correlation coefficients for the change in weight $z$ scores for each interval, as well as birth weight and BMI at 7 years of age. The purpose of Table 2 is to demonstrate that other than birth weight and BMI at 7 years of age, changes in growth percentiles were not strongly correlated between intervals. In other words, change in size percentile in 1 interval was not predictive of change in size percentile in other intervals. This is important because it then allows us to use the growth data for every interval together in logistic regression analysis, without the risk for collinearity.

In logistic regression analysis, we used birth weight to predict high BP at 7 years of age in an unadjusted model.

---

**TABLE 1** Maternal Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean (SD or %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>24.5 (6.1)</td>
</tr>
<tr>
<td>Prepregnancy BMI, kg/m²</td>
<td>22.9 (4.3)</td>
</tr>
<tr>
<td>Education, y</td>
<td>10.9 (2.4)</td>
</tr>
<tr>
<td>White race</td>
<td>52.3</td>
</tr>
<tr>
<td>Smoking</td>
<td>46.7</td>
</tr>
<tr>
<td>Anemia (hematocrit &lt;30%)</td>
<td>14.2</td>
</tr>
<tr>
<td>Type 1 diabetes &gt; 5 y</td>
<td>0.5</td>
</tr>
<tr>
<td>Poverty</td>
<td>54.6</td>
</tr>
<tr>
<td>Toxemia</td>
<td>2.8</td>
</tr>
<tr>
<td>Pregnancy-induced hypertension</td>
<td>3.4</td>
</tr>
<tr>
<td>Chronic hypertension</td>
<td>5.4</td>
</tr>
</tbody>
</table>

**TABLE 2** Correlation Matrix for Birth Weight and Change in Weight $z$ Scores During Childhood

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Birth Weight</th>
<th>Birth Weight to 4-mo $\Delta z$ Score</th>
<th>4-mo to 1-y $\Delta z$ Score</th>
<th>1- to 4-y $\Delta z$ Score</th>
<th>4- to 7-y $\Delta z$ Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birth weight to 4-mo $\Delta z$ score</td>
<td>-0.49</td>
<td>-0.19</td>
<td>-0.28</td>
<td>-0.18</td>
<td>-0.04</td>
</tr>
<tr>
<td>4-mo to $z$ score</td>
<td>-0.03</td>
<td>-0.11</td>
<td>-0.18</td>
<td>-0.04</td>
<td>-0.04</td>
</tr>
<tr>
<td>1- to 4-y $\Delta z$ score</td>
<td>-0.14</td>
<td>-0.04</td>
<td>-0.05</td>
<td>-0.04</td>
<td>0.38</td>
</tr>
<tr>
<td>4- to 7-y $\Delta z$ score</td>
<td>0.19</td>
<td>0.10</td>
<td>0.06</td>
<td>0.38</td>
<td>0.36</td>
</tr>
</tbody>
</table>

All correlations are significant at the $P < .01$ level (2-tailed).
Each 1-kg increase in birth weight increased the odds for high SBP by 1.06 (0.98–1.14; \( P = .16 \)) and high DBP by 1.11 (1.03–1.21; \( P = .007 \)) and widened PP by 0.95 (0.88–1.03; \( P = .23 \)). In multivariable logistic regression, we included race and change in weight z scores throughout childhood as independent variables to predict BP. Race was added to the model because in bivariate analysis, we found that of maternal characteristics listed in Table 1, only race had a significant influence on childhood BP (data not shown). Table 3 shows regression models that predicted high SBP, DBP, and widened PP, respectively. Each 1-kg increase in birth weight adjusted for race and change in weight z scores more than doubles the risk for high SBP at 7 years of age (odds ratio [OR]: 2.19; 95% confidence interval [CI]: 1.92–2.49; \( P < .001 \)). White race increases the odds by 1.51 (95% CI: 1.36–1.66; \( P < .001 \)). A 1-U increase in weight z score between 2 ages increases the risk for high SBP anywhere between 1.65 to 1.94 times, depending on the interval (\( P < .001 \)). After adjustment for race, the odds for high DBP for each 1-kg increase in birth weight were increased by 1.82 (95% CI: 1.59–2.08; \( P < .001 \)), and white race also increased the odds, albeit modestly (OR: 1.28; 95% CI: 1.16–1.42; \( P < .001 \)). Change in weight z scores had a positive but modest predictive effect on high DBP (OR range: 1.43–1.56; \( P < .001 \)). Birth weight had the smallest effect on PP, with an OR of 1.22 in predicting a widened PP (95% CI: 1.06–1.39; \( P = .005 \)). White race increased the odds by 1.46 (95% CI: 1.06–1.39), and change in weight z scores had a small positive association with widened PP (OR: 1.20–1.33; all \( P < .001 \)). Within each of the 3 models in Table 3, the ORs for high BP per unit change in the 4 growth intervals were not statistically different from each other.

We tested the interactions between birth weight and catch-up growth by including interaction terms in regression models to predict high SBP, DBP, and PP. The interaction terms between birth weight and change in weight z score for each interval each were run in a separate model, with birth weight, race, and the 4 changes in weight z scores as the other independent variables. None of the interaction terms was found to be statistically significant (therefore, we chose not include the raw data in this article). The lack of interaction between birth weight and changes in weight z score suggests that the increased risk for high BP that is caused by rapid postnatal growth occurs to a similar degree regardless of size at birth. To illustrate this point, we stratified our analysis by size at birth (SGA, AGA, and LGA) and applied the previously described logistic regression model to predict high SBP (Table 4). We chose SBP because it is the most commonly cited measure of BP in the literature on birth weight and fetal programming. As demonstrated in Table 4, we found that an increase of 1 U in weight z score during any of the defined intervals increased the odds for high SBP to a similar degree in all 3 groups of infants. A 1-kg increase in birth weight increases the odds for high SBP by 3.53 in infants who are SGA (95% CI: 1.20–10.34; \( P = .02 \)).

### Table 3

<table>
<thead>
<tr>
<th>Risk Variable</th>
<th>ORa</th>
<th>95% CI</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable: high SBP at age 7 y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight, kg</td>
<td>2.19</td>
<td>1.92–2.49</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>White race</td>
<td>1.51</td>
<td>1.36–1.66</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, birth to 4 mo</td>
<td>1.65</td>
<td>1.54–1.75</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 4 mo to 1 y</td>
<td>1.79</td>
<td>1.66–1.93</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 1 to 4 y</td>
<td>1.71</td>
<td>1.61–1.80</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 4 to 7 y</td>
<td>1.94</td>
<td>1.81–2.08</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Dependent variable: high DBP at age 7 y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight, kg</td>
<td>1.82</td>
<td>1.59–2.08</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>White race</td>
<td>1.28</td>
<td>1.16–1.42</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, birth to 4 mo</td>
<td>1.43</td>
<td>1.34–1.52</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 4 mo to 1 y</td>
<td>1.44</td>
<td>1.34–1.56</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 1 to 4 y</td>
<td>1.45</td>
<td>1.37–1.53</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 4 to 7 y</td>
<td>1.56</td>
<td>1.45–1.67</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Dependent variable: widened PP at age 7 y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight, kg</td>
<td>1.22</td>
<td>1.06–1.39</td>
<td>.005</td>
</tr>
<tr>
<td>White race</td>
<td>1.46</td>
<td>1.32–1.61</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, birth to 4 mo</td>
<td>1.20</td>
<td>1.13–1.28</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 4 mo to 1 y</td>
<td>1.31</td>
<td>1.21–1.41</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 1 to 4 y</td>
<td>1.22</td>
<td>1.15–1.29</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 4 to 7 y</td>
<td>1.33</td>
<td>1.23–1.43</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

* ORs reflect risk adjusted for other variables in model: birth weight, race, and change in SBP z scores.

### Table 4

<table>
<thead>
<tr>
<th>Risk Variable</th>
<th>ORa</th>
<th>95% CI</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infants who were SGA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight, kg</td>
<td>3.53</td>
<td>1.20–10.34</td>
<td>.022</td>
</tr>
<tr>
<td>White race</td>
<td>1.97</td>
<td>1.33–2.90</td>
<td>.001</td>
</tr>
<tr>
<td>Weight gain, birth to 4 mo</td>
<td>1.65</td>
<td>1.26–2.15</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 4 mo to 1 y</td>
<td>1.90</td>
<td>1.39–2.59</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 1 to 4 y</td>
<td>1.61</td>
<td>1.29–2.00</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 4 to 7 y</td>
<td>1.67</td>
<td>1.27–2.21</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Infants who were AGA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight, kg</td>
<td>2.16</td>
<td>1.82–2.55</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>White race</td>
<td>1.52</td>
<td>1.36–1.68</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, birth to 4 mo</td>
<td>1.65</td>
<td>1.54–1.76</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 4 mo to 1 y</td>
<td>1.77</td>
<td>1.63–1.91</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 1 to 4 y</td>
<td>1.74</td>
<td>1.63–1.84</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 4 to 7 y</td>
<td>1.97</td>
<td>1.83–2.12</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Infants who were LGA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight, kg</td>
<td>3.31</td>
<td>1.30–8.40</td>
<td>.012</td>
</tr>
<tr>
<td>White race</td>
<td>0.99</td>
<td>0.64–1.51</td>
<td>.956</td>
</tr>
<tr>
<td>Weight gain, birth to 4 mo</td>
<td>1.66</td>
<td>1.30–2.12</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 4 mo to 1 y</td>
<td>2.02</td>
<td>1.53–2.66</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 1 to 4 y</td>
<td>1.44</td>
<td>1.17–1.78</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weight gain, 4 to 7 y</td>
<td>1.89</td>
<td>1.48–2.42</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

* ORs reflect risk adjusted for other variables in model: birth weight, race, and change in SBP z scores.

a OR for 1-U change in weight z scores during interval.
2.16 in infants who are AGA (95% CI: 1.82–2.55; \( P < .001 \)), and 3.31 in infants who are LGA (95% CI: 1.30–8.40; \( P = .01 \)). White race increases the odds for high SBP in infants who are AGA and SGA but not in infants who are LGA (OR: 0.99; \( P = .96 \)).

**DISCUSSION**

This analysis of the CPP has led us to 4 conclusions. First, infants who are SGA are not at higher risk for high BP at 7 years of age in this cohort. Second, children who cross weight percentiles upward are at increased risk for high BP in early childhood. During each childhood growth interval, there was a modest increase in the odds for high SBP at 7 years of age for each SD of increase in relative size during that interval. Third, the timing of the upward crossing of weight percentiles does not modify the risk: the increase in ORs for high SBP was similar for each of the 4 intervals that we studied, leading us to conclude that increasing growth percentiles during any period of early childhood increases the risk for high BP. Fourth, there is no interaction between size at birth and postnatal growth in predicting high BP in early childhood, on the basis of our observation that the magnitude of the effect of weight gain on BP does not depend on size at birth.

This analysis was conducted on prospectively collected data from a large biracial cohort of pregnant women and their offspring. To our knowledge, this is the largest study of its kind conducted on a US data set, and the very large sample size makes it 1 of the few American data sets appropriate for use in studying the effect of birth size and subsequent catch-up growth in childhood on BP. The repeated weight measurements during the study follow-up made it possible to calculate change in weight \( z \) scores for 4 intervals in early childhood. This allowed us to avoid the bias that was introduced by the fact that a small neonate requires less absolute weight gain than a large neonate to maintain growth percentile. The advantage to quantifying catch-up growth by change in weight \( z \) score is the effective independence of each measure from the starting size (birth weight) and each of the other growth measures. This allowed us to consider all 4 interval measures, along with birth weight, simultaneously in our regression models.

Despite the unique nature of this study, there were several limitations to the data. Birth weight was used in this analysis as a proxy measure of fetal growth restriction, a sign that the in utero environment may have been compromised in a way that would also lead to changes in fetal physiology (programming). However, the use of birth weight <10th percentile for gestational age as a screening tool for growth restriction is less than ideal because some infants who are constitutionally small but not growth restricted will be inadvertently included in this group. Our data, like most other studies of IUGR, are subject to this limitation. Future studies using serial fetal ultrasonography or customized fetal growth curves to identify IUGR instead of birth weight are warranted.

In addition, the CPP used last menstrual period for pregnancy dating. Gross inaccuracies in dating were excluded from analysis by the removal of outlier data for gestational age and birth weight, but more subtle inaccuracies may have led to the misclassification of infants whose birth weights were borderline SGA. A previously published report demonstrated that when gestational age dating on the basis of last menstrual period indicates a pregnancy is at term, ultrasound is usually within 1 week of the menstrual estimate. However, menstrual dates are frequently in error when they indicate that the infant is either preterm or postterm. Because our study included only infants with term gestation, we do not believe that the use of menstrual dating in the CPP introduces any significant error into our analysis.

A third caveat is that the CPP followed children until 7 years, an age at which only very small differences in BP are detected between individuals who later become normotensive versus hypertensive adults. These differences are expected to be amplified as the children grow older. Identifying significant differences within the tight range of BPs in 7-year-old children is compounded by the possibility of measurement error in the CPP study, which required only 1 BP measurement at the 7-year clinic visit. We have no reason to believe that this error was systematic or that it affected normal and abnormal BP differentially. Therefore, we do not expect the chance for random error in BP measurement in the CPP study to affect the results of our analysis.

Contrary to many published studies, we found that birth weight and BP later in life are positively associated. The most likely reason for this discrepancy is our decision not to include current size (BMI) in the model, as many previous studies have done. We elected not to include BMI at 7 years of age in our primary model because of what is commonly known in the statistical literature as the “reversal paradox.” The reversal paradox refers to the seeming reversal of a statistical association (positive to negative, or vice versa) between 2 variables when a third etiologic variable is introduced into the regression model. If the third variable is actually on the causal pathway between the first 2 variables, then the inclusion of that third variable may invert the association between the other 2. As several authors have noted, current BMI may be on the causal pathway between birth size and hypertension; therefore, the inclusion of BMI in models of birth weight that predict BP may actually reverse the seeming statistical association between birth weight and BP. In addition, it has been pointed out that BMI not only is positively related to birth weight but also is a much more powerful predictor of hypertension. Therefore, controlling for BMI in the model would cancel out the positive effects of birth
weight on BMI as well as BP.29 We therefore chose not to include BMI in our regression model.

Our finding of a positive direction of association between birth weight and BP is comparable to other published data that do not include a measure of current weight in regression analysis.29 What is more difficult to compare is the positive influence of postnatal growth on BP that we report here. The varied methods of quantifying catch-up growth in the literature are not easily comparable to one another. Our method of using change in z scores for weight has been used in at least 2 other published studies. The first reported that in a population of 346 British men and women, an increase in weight z score between birth and 1 year or between 1 year and 5 years did not increase SBP or DBP at 22 years of age (with or without adjustment for adult BMI).30 Our results may have differed for 1 of several reasons: we used tighter age intervals, examined childhood BP, and had a population that included both white and black Americans. The second study reported that in a population of 749 Brazilian adolescents, an increase of 1 U of weight z score per year resulted in an increase in SBP of 0.37 mm Hg at 15 years of age, and weight gain in infancy, childhood, or adolescence had the same implications on BP.31 This positive and similar effect of weight gain at different intervals of childhood growth on BP is comparable to our results.

CONCLUSIONS
On the basis of our data, an increase in growth percentile during the first 7 years of life puts a child at increased risk for high BP in early childhood. Because BP tends to track over a lifetime, this rapid growth has implications for adult hypertension.32,33 There is a continuum of risk for hypertension across the birth weight spectrum, so we must carefully consider the wisdom of encouraging rapid weight gain in any of our patients, including but not limited to infants with IUGR and preterm infants, in light of their future health. Future research should attempt to determine in which children catch-up growth is actually “excess growth” and which infants and children would benefit by maintaining their growth percentiles throughout childhood and beyond. Clearly, additional study is important for both pediatric and adult health.

ACKNOWLEDGMENTS
Drs Hemachandra, Howards, and Klebanoff are supported by the intramural research program of the National Institute of Child Health and Human Development, National Institutes of Health. Dr Furth is supported by National Institute of Diabetes and Digestive and Kidney Diseases grant U01DK66174.

REFERENCES


Firearm Ownership and Storage Patterns Among Families With Children Who Receive Well-Child Care in Pediatric Offices

Robert H. DuRant, PhD, Shari Barkin, MD, MSHS, Joseph A. Craig, MD, Victoria A. Weiley, MIS, Edward H. Ip, PhD, Richard C. Wasserman, MD, MPH

Departments of Pediatrics, Social Science and Health Policy, and Biostatistical Sciences, Wake Forest University School of Medicine, Winston-Salem, North Carolina; Rocky Mountain Youth Clinics, Denver, Colorado; Pediatric Research in Office Settings, American Academy of Pediatrics, Elk Grove Village, Illinois; Department of Pediatrics, University of Vermont, Burlington, Vermont

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

ABSTRACT

OBJECTIVE. In this study we examined firearm storage patterns and their associations in a diverse sample of families who attended pediatric practices from both rural and nonrural areas across the United States.

METHODS. Parents who brought their children who were aged 2 to 11 years (N = 3745) to 96 Pediatric Research in Office Settings practices from 45 states, Canada, and Puerto Rico participated in an office-based survey before a well-child examination. The survey measured demographic variables; family history of guns in the home; and firearm types, storage behaviors, and ownership.

RESULTS. Twenty-three percent of families reported firearm ownership. The majority (60%) of respondents reported making firearm storage decisions. Only one third of firearm owners reported safe firearm storage. Gun type owned was associated with storage habits, with long-gun owners storing their gun in places other than locked cabinets but with ammunition separate from guns and handgun users more likely to store guns loaded and to use gun locks. In a multivariate analysis, not being raised with a firearm was associated with safe storage behaviors. Families who had children aged 2 to 5 years and owned long guns were more likely to store their guns safely than families with older children.

CONCLUSIONS. Few families reported safe firearm storage. Storage patterns are most influenced by firearm type(s) owned, family socialization with guns, and the age of the child. Primary care providers need to understand better not only whether firearms are in the home but also which types are present and whether parents were raised in homes with guns.
There are between 192 and 200 million privately owned guns in the United States; ~33% to 41% of all US homes report at least 1 gun in the home.1–4 Gun ownership varies nationally from 5.2% of homes in the District of Columbia to 62.8% in Wyoming.2 Among adults with children who are younger than 18 years, the prevalence of loaded household firearms ranged from 1% in Massachusetts to 13.4% in Alabama.2 More than 22 million children (35% of homes with children) live in homes in which respondents reported the presence of a firearm.3 Among homes with children and firearms, between 41.5% and 43% had at least 1 unlocked firearm,3 which compared with 20% of all gun-owning households that stored guns unlocked and loaded.5

Although the findings from the research vary, these data indicate that substantial numbers of pediatricians’ patients either live in or visit a home where a gun is present.4–7 There is strong evidence attesting to the magnitude and the nature of the threat that is posed by the prevalence of firearms, especially when stored unsafely.3,8–10 The accessibility of guns in the home has been associated with an increased risk for suicide and unintentional injury.11,12 A recent case-control study of youth suicide and unintentional injury with firearms showed that the individual practices of storing guns (1) locked, (2) unloaded, (3) with ammunition locked, and (4) with ammunition in a separate location from the gun each significantly reduced the risk to children for a firearm-related injury.8 If children and adolescents are living in a home with a firearm, then proper gun storage is critical to lessening their risk for injury or death.7,11–16 Organizations that support gun ownership and those that are involved in child advocacy recommend safe storage practices.17,18 Surveys of gun-owning families with children, however, have revealed that from 43% to 95% stored at least 1 gun unlocked and 9% to 20% stored at least 1 gun loaded.2,3,7,14–16 Differences in firearm storage have been cited according to geographic region. In Connor’s18 recent study of random samples of urban and rural adults, 31% of rural households, as compared with only ~13% of urban households, reported firearm ownership. Respondents with children in the home reported significantly lower gun ownership (20%) than respondents without children (29%). However, safe firearm storage (locked or locked up and separate from ammunition) did not differ significantly on the basis of the presence of children in the home after controlling for rural versus urban setting; safe storage patterns did not differ from rural versus urban families. However, Nordstrom et al19 found that although the prevalence of firearms in the home was higher in rural areas than in nonrural areas, the prevalence of loaded, unlocked guns was actually lower in rural areas than in urban areas. Further complicating the issue are the studies reporting that >1 firearm is more likely to be found in homes in rural areas than in urban regions,3,14 creating a situation of children’s increased exposure to potentially unsafely stored firearms. Even so, Nance et al20 found that rates of serious firearm injuries among children and adolescents were 10-fold higher in urban than in nonurban regions.

Rural-urban differences in gun ownership and storage practices reflect long established cultural differences in the types and the number of guns owned and the reasons for firearms ownership.5,6,14,20,21 For example, in rural areas, many gun owners possess firearms for recreational purposes, in contrast to urban areas, where many guns are owned for protection.3,14,20–22 In Baxley and Miller’s21 study of rural parents, owners of guns for recreation were more likely to store guns locked and unloaded (70%) than parents who owned guns for protection (52%). Other factors that were found to be associated with safe firearm storage in various populations include having a 4-year college education, total family income of $65 000 per year or more,16 owning handguns in contrast to long guns,6,19 children’s knowing the storage location of the firearms in the home and having handled firearms,21 living in a rural town rather than a farm household, not taking a gun safety course, and the absence of a history of drug or alcohol abuse in the family.19

Although considerable research has been done to identify factors that are associated with firearm storage practices of parents who seek primary health care for their children, there is still a lack of consensus on what are the most important risk factors for unsafe firearm storage. Part of the reason is that definitions for rural, urban, nonrural, and nonurban have varied across studies or have not been defined at all.3,6,12,14,19–22 As expected, risk factors that have been found to be specific to geographic locations have also varied. In addition, many previous studies have lacked racial/ethnic diversity, differed in how safe firearm storage was defined, been limited in geographic scope, and varied in whether the respondent was responsible for making firearm storage decisions in the home.3,7,14

For development of more effective provider counseling approaches about gun safety, variables that are associated with firearm ownership and storage practices need to be specified. In this study we examined differences between gun-owning and non–gun-owning families from a diverse sample of pediatric practices in both rural and nonrural areas across the United States. We examined storage patterns and their associations reported by firearm owners.

METHODS

Pediatric Research in Office Settings (PROS), the practice-based research network of the American Academy of Pediatrics (AAP), conducted this study as part of a group randomized, interventional study to reduce violence by teaching families skills to decrease media use, increase the use of noncorporal discipline, and promote...
safe firearm storage. Using standard methods for all PROS studies, all PROS practices were sent information and an invitation to participate in the study. State coordinators of PROS were asked to identify and recruit practices for the study. State coordinators were particularly helpful in recruiting pediatric practices that serve large populations of minority racial/ethnic children. Parents and guardians who brought their children to 138 PROS practitioners in 96 practices from 45 states, Canada, and Puerto Rico participated in an office-based survey before the well-child visit for children who were 2 to 11 years of age (N = 3745) (see “Acknowledgments” for a list of participating practices). Latino and rural parents were oversampled at the practice level to ensure adequate sample sizes of these subgroups for analyses. Enrollment was limited to 1 child per family. Institutional review board approval was obtained from the Wake Forest University School of Medicine and the AAP.

Parents who brought their children to PROS practices for well-child visits were invited by trained research coordinators in each office to participate in a study titled Safety Check. When the parent agreed verbally to participate, the parent provided written informed consent and completed a previst questionnaire while in the waiting room or the examination room. Only the legal guardian/parent who reported spending the majority of time with the child when at home served as the respondent. The questionnaires were kept confidential, containing only a study number that could link the previst questionnaire to 2 postvisit telephone surveys. Parents put their surveys in a sealed envelope after completing it, and the participating pediatricians did not have access to this survey. Each pediatrician enrolled parents/legal guardians until 30 parents were enrolled per provider. Of 4499 parents, 266 were determined to be ineligible and 485 (11.5%) of 4230 refused to participate in the study. Of the 3745 completed pretest surveys, 204 were completed in Spanish. The data for this article were derived from baseline surveys that were gathered before the patient visit.

Survey Instrument

The distinction between rural and nonrural practice settings was determined by asking the health care provider the question, “How would you describe the area in which your practice is located?” with response categories “urban, inner city,” “urban, not inner city,” “suburban,” and “rural.” This method has been used by PROS since 1993 to classify practice demographics. From 1993 to 1997, US census metropolitan statistical area population categories were also used, but these were discontinued in 1998 because of the redundancy of the classifications. We considered using rural/urban community areas (RUCAs) zip code approximations to classify rural-urban practice setting. On the basis of recommendations by Hart et al,25 we did not use this method because of the disconnection between health data from the local level and RUCA codes and the lack of stability of RUCA codes over time in general. Because firearm storage patterns did not differ among urban inner city, urban not-inner city, and suburban, we collapsed these into a single subgroup that we termed nonrural. Parents/legal guardians completed questions that included demographic variables (eg, age of child, number of children in the home, race/ethnicity of the child, parental home structure, maternal education), firearm ownership, firearm storage behaviors, parental firearm concern, firearm owner (responsible for firearm storage decisions), and parental history of guns in the home and owned for protection (firearm family socialization).

Firearm storage–related behavior questions included the following: “Are any guns stored or hidden in a place other than a locked cabinet or gun safe?” “Are all guns stored with a gunlock on them?” “Are bullets stored separate from all guns?” These responses were rated as “yes,” “no,” or “don’t know.” To gauge parental concern about children’s exposure to firearms, parents responded “yes” or “no” to the question, “I worry that if this child found a gun, [he or she] would play with it.” Questions were also asked about the primary caregiver’s own childhood experience with firearms (firearm family socialization): “I was raised in a home with guns in it,” and, “My family owned a gun for protection.”

Statistical Analysis

Descriptive summaries were obtained to provide a comparison between the characteristics of gun-owning households (n = 872) and non–gun-owning households (n = 2800). We then compared gun ownership and storage patterns by rural versus nonrural status. Firearm storage patterns were compared across 3 groups: (1) handgun owner only, (2) long-gun owner only (rifle or shotgun), or (3) combination gun type owner (both long guns and handguns). We then identified the subsample of gun-owning respondents who reported that they were the firearm storage decision-maker (n = 516) by asking the question, “In your home, who makes decisions most of the time about gun storage?” For these, we conducted a bivariate analysis of firearm storage variables related to handgun only, long gun only, and combination ownership. Finally, we performed a multivariate logistic regression analysis of the factors from the literature in that would be expected to affect safe firearm storage practices. The 3 survey items that were used to form the dependent variable of safe storage included all guns stored in locked cabinet or safe, all guns stored with gunlocks on them, and bullets stored separate from all guns. The list of independent variables included (1) gun type, with long gun only or combination ownership compared with handgun only ownership; (2) firearm family socialization; and (3) demographics. The demographic variables included rural
versus nonrural, child’s race/ethnicity, child’s age, and maternal education level. Race/ethnicity categories included white, black, Latino, multirace, and “other” (when parents checked multiple ethnicities, the child was coded as “multirace”). Household income was found to have a high correlation with maternal education; therefore, income was not included in the final model.

RESULTS

The study sample was drawn from 96 practices from 45 states, Canada, and Puerto Rico and oversampled practices with Latino children (19% vs 14%) and families who lived in rural areas (37% vs 21%). This sample approximated national census data regarding black and white children and maternal education level. Ninety percent of the parents who completed the previsit survey were female.

For the complete sample (N = 3745), 23.2% of families reported gun ownership. The rates of firearm ownership were lower in nonrural regions (19.6%) compared with rural regions (34.4%). Table 1 displays the demographic variables of the overall sample and compares the demographic variables of firearm owners with non–firearm owners. White families were more likely to own firearms than families from other race/ethnic groups. Firearm ownership was lowest in families in which parental education was less than a high school graduate. Gun ownership was highest in families with 2 adults in the home when compared with any other family configuration. Also, families with 3 or more adolescents in the home were less likely to own firearms.

Among gun owners (n = 872), families who lived in rural areas were more likely than families who lived in nonrural areas to have only long guns in the home, whereas families in nonrural areas were more likely to own only handguns (Table 2). Both rural and nonrural families exhibited the same degree of combination firearm ownership. Among gun owners, families who lived in rural areas were more likely than those in nonrural areas to have been raised in homes in which their family owned 1 or more guns. However, when only parents who made decisions about how guns were stored in

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Demographic Variables Comparing the Percentage of Gun-Owning Households With Non–Gun-Owning Households</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
<td>Overall Sample</td>
</tr>
<tr>
<td></td>
<td>%</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>61.8</td>
</tr>
<tr>
<td>Black</td>
<td>11.9</td>
</tr>
<tr>
<td>Hispanic</td>
<td>18.7</td>
</tr>
<tr>
<td>Multirace</td>
<td>3.6</td>
</tr>
<tr>
<td>Other</td>
<td>4.1</td>
</tr>
<tr>
<td>Child’s age, y</td>
<td></td>
</tr>
<tr>
<td>2–5</td>
<td>52.9</td>
</tr>
<tr>
<td>6–11</td>
<td>47.1</td>
</tr>
<tr>
<td>Parental education</td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>9.7</td>
</tr>
<tr>
<td>High school graduate and associate degree</td>
<td>55.5</td>
</tr>
<tr>
<td>College graduate and above</td>
<td>34.2</td>
</tr>
<tr>
<td>No. of children younger than 12 y</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>4.6</td>
</tr>
<tr>
<td>1</td>
<td>30.1</td>
</tr>
<tr>
<td>2</td>
<td>41.0</td>
</tr>
<tr>
<td>≥3</td>
<td>24.3</td>
</tr>
<tr>
<td>No. of children between 12 and 18 y</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>75.9</td>
</tr>
<tr>
<td>1</td>
<td>16.2</td>
</tr>
<tr>
<td>2</td>
<td>5.9</td>
</tr>
<tr>
<td>≥3</td>
<td>2.0</td>
</tr>
<tr>
<td>No. of adults at home</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>14.6</td>
</tr>
<tr>
<td>2</td>
<td>76.5</td>
</tr>
<tr>
<td>≥3</td>
<td>8.9</td>
</tr>
<tr>
<td>Family income, $</td>
<td></td>
</tr>
<tr>
<td>&lt;39 999</td>
<td>40.3</td>
</tr>
<tr>
<td>40 000–79 999</td>
<td>33.0</td>
</tr>
<tr>
<td>≥80 000</td>
<td>26.7</td>
</tr>
</tbody>
</table>

Approximately 2% of the respondents said that they did not know whether they own a gun; therefore, the percentages for gun owners and non–gun owners may not add up to exactly 100%. P-values are based on χ²-tests.
their home were analyzed, the significance level of this relationship was reduced from $P = .02$ to .07.

Among gun-owning families, the mother was the respondent in 87.5% of the cases. In only 9.8% of the cases did she report that she was the sole decision-maker about gun storage in the home, but in 47.9% of the cases, she made storage decisions jointly with others in the home. Among the 8.5% of fathers who were the respondent, 73.2% reported that they were the sole decision-maker about gun storage, but the rest (26.8%) reported that they made decisions jointly with someone else in the family. Decision-makers were more likely to report that guns were stored safely (35.9%) than non–decision-makers (25.2%; $P = .001$). Because of this, we examined characteristics of gun storage among decision-makers only to determine gun storage practices.

Examining firearm decision-makers only ($n = 516$), the types of guns owned were significantly associated with the gun storage practices of families (Table 3). Families who owned only long guns were more likely than families who owned only handguns or a combination of gun types to store their guns in places other than locked cabinets. The use of gun locks on all of the guns in the home was highest among handgun-only owners, followed by long gun–only owners, and then combination gun owners. In contrast, the practice of storing bullets and shells separately from guns was highest among long-gun owners, followed by combination gun owners. Handgun-only owners were the most likely parents to store guns and bullets together. Families who owned both long guns and handguns were the most likely to have been raised in a family with guns, followed by long gun–only owners. Families who were combination gun owners were also more likely to report that they owned a gun for protection. Handgun-only owners were the most likely to believe that it was very important to use gun locks with guns. When these same relationships were examined for all of the respondents who reported having a firearm in the home (regardless of whether they were decision-makers), the relationships did not change.

When storing all guns in the home safely (as defined above) was examined using multiple logistic regression limited to the subsample of families who reported being the firearm storage decision-maker, safe firearm storage was similar among white (36.5%), black (37.9%), and

### Table 2: Rural Versus Nonrural Differences Among Gun-Owning Households ($N = 872$)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Rural ($N = 320$)</th>
<th>Nonrural ($N = 552$)</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only long guns in the house</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>82.8</td>
<td>61.2</td>
<td>0.001</td>
</tr>
<tr>
<td>No</td>
<td>17.2</td>
<td>38.8</td>
<td></td>
</tr>
<tr>
<td>Only handguns in the house</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>13.8</td>
<td>35.6</td>
<td>0.001</td>
</tr>
<tr>
<td>No</td>
<td>86.2</td>
<td>64.4</td>
<td></td>
</tr>
<tr>
<td>Combination of long guns and handguns in the house</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>43.1</td>
<td>37.9</td>
<td>0.13</td>
</tr>
<tr>
<td>No</td>
<td>56.9</td>
<td>62.1</td>
<td></td>
</tr>
<tr>
<td>Raised in home with guns</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>64.7</td>
<td>53.8</td>
<td>0.02</td>
</tr>
<tr>
<td>No</td>
<td>30.9</td>
<td>41.5</td>
<td></td>
</tr>
<tr>
<td>Don’t know</td>
<td>3.4</td>
<td>3.4</td>
<td></td>
</tr>
<tr>
<td>Stores guns safely</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>34.2</td>
<td>36.8</td>
<td>0.20</td>
</tr>
<tr>
<td>No</td>
<td>65.8</td>
<td>63.2</td>
<td></td>
</tr>
<tr>
<td>Family owned gun for protection</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>30.3</td>
<td>30.4</td>
<td>0.79</td>
</tr>
<tr>
<td>No</td>
<td>65.0</td>
<td>63.6</td>
<td></td>
</tr>
<tr>
<td>Don’t know</td>
<td>4.1</td>
<td>4.7</td>
<td></td>
</tr>
<tr>
<td>Makes decision about gun storage</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Me</td>
<td>15.6</td>
<td>14.5</td>
<td>0.60</td>
</tr>
<tr>
<td>Other adults</td>
<td>36.3</td>
<td>37.3</td>
<td></td>
</tr>
<tr>
<td>Me and other</td>
<td>44.7</td>
<td>44.0</td>
<td></td>
</tr>
<tr>
<td>Important to use gunlocks</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not important</td>
<td>3.1</td>
<td>3.4</td>
<td>0.79</td>
</tr>
<tr>
<td>Somewhat important</td>
<td>17.8</td>
<td>11.6</td>
<td></td>
</tr>
<tr>
<td>Very important</td>
<td>73.8</td>
<td>81.3</td>
<td></td>
</tr>
<tr>
<td>Worry that if child found gun would play with it</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>32.1</td>
<td>30.9</td>
<td>0.06</td>
</tr>
<tr>
<td>No</td>
<td>64.4</td>
<td>59.2</td>
<td></td>
</tr>
</tbody>
</table>

Because of missing values, the percentages within a cell may not add up to exactly 100%. $P$ values are based on $\chi^2$ tests.
Latino (43.2%) families (Table 4). Multirace families were less likely to store guns safely. Parents who were raised in families where guns were not owned (41.7%) were 66% more likely to engage in safe gun storage than were parents raised in gun-owning families (32.3%). When we examined interaction effects in the model, an age \times gun type interaction was found (odds ratio: 5.26; 95% confidence interval: 1.65–16.72). When compared with families with children who were 6 to 11 years of age, families who had children who were 2 to 5 years of age and owned long guns were more likely to store guns safely. No other variables were significantly associated with safe gun storage after these variables were taken into account. Although the descriptive data indicate that gun ownership was lowest in families in which the respondent had less than a high school education, once this was controlled for in a multivariate model, this finding did not remain significant.

**DISCUSSION**

In the United States, many children and adolescents are living in homes where firearms are stored.2–7 In this study of parents across the United States who brought their children to pediatric offices, 23.3% of households reported the presence of a firearm, which is lower than has been reported in other studies. In 2 population-based studies, 32.6% (range: 5.2%–62.8%) of all families and 35% of families with children3 reported guns in the home. However, in a large study of mostly Hispanic parents of children in 1 pediatric clinic, gun ownership was only 7.8%.16 In our study, firearm ownership was highest in rural families and in families that were white, had 2 or fewer adolescents in the home, had 2 adults in the home, and had a total family income of $40 000 per year or more. Also, 19% of our sample were Latino and

### Table 3

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Handguns Only (N = 89)</th>
<th>Long Guns Only (N = 126)</th>
<th>Combination (N = 202)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>% n</td>
<td>% n</td>
<td>% n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guns hidden in place other than locked cabinet</td>
<td>31.8 28</td>
<td>37.5 47</td>
<td>25.7 52</td>
<td>.04</td>
</tr>
<tr>
<td>All guns stored with lock</td>
<td>75.6 67</td>
<td>63.2 80</td>
<td>50.5 102</td>
<td>.0002</td>
</tr>
<tr>
<td>Bullets stored separately</td>
<td>70.8 63</td>
<td>89.4 113</td>
<td>79.2 160</td>
<td>.0003</td>
</tr>
<tr>
<td>Raised in home with guns</td>
<td>44.9 40</td>
<td>58.3 73</td>
<td>72.3 146</td>
<td>.0004</td>
</tr>
<tr>
<td>Family owned gun for protection</td>
<td>29.2 26</td>
<td>24.5 31</td>
<td>44.1 89</td>
<td>.0002</td>
</tr>
<tr>
<td>Worry that if child found gun would play with it</td>
<td>42.7 38</td>
<td>32.9 41</td>
<td>34.2 69</td>
<td>.04</td>
</tr>
<tr>
<td>Because of missing values, the percentages within a cell may not add up to exactly 100%. P values are based on χ² tests.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 4

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR</th>
<th>95% CI</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child’s age</td>
<td>0.90</td>
<td>0.60–1.33</td>
<td>.59</td>
</tr>
<tr>
<td>Maternal education</td>
<td>0.93</td>
<td>0.64–1.35</td>
<td>.71</td>
</tr>
<tr>
<td>Child’s race/ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.77</td>
<td>0.32–1.83</td>
<td>.55</td>
</tr>
<tr>
<td>Latino</td>
<td>1.05</td>
<td>0.50–2.21</td>
<td>.89</td>
</tr>
<tr>
<td>Multirace</td>
<td>0.13</td>
<td>0.02–0.99</td>
<td>.05</td>
</tr>
<tr>
<td>Other</td>
<td>0.11</td>
<td>0.01–0.84</td>
<td>.03</td>
</tr>
<tr>
<td>Gun type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combination</td>
<td>0.83</td>
<td>0.45–1.53</td>
<td>.55</td>
</tr>
<tr>
<td>Long gun</td>
<td>0.99</td>
<td>0.55–1.78</td>
<td>.96</td>
</tr>
<tr>
<td>Nonrural</td>
<td>1.19</td>
<td>0.77–1.84</td>
<td>.43</td>
</tr>
<tr>
<td>Not raised in home with gun</td>
<td>1.66</td>
<td>1.09–2.53</td>
<td>.02</td>
</tr>
</tbody>
</table>

The dependent variable was safe storage of firearms, defined as having all guns in a locked location, gun locks used, and bullets/shells stored separately from guns. The reference categories for the independent variables included younger child aged 2 to 5 years, less than a college education, white, handgun-only ownership, rural site, and raised in a home with firearms. OR indicates odds ratio; CI, confidence interval.
12% were black. Because these 2 populations report firearm ownership to a much lesser degree than white individuals, it is possible that this resulted in fewer firearm owners in our total sample.

Families in rural areas were more likely to report firearm ownership; however, gun ownership among rural families was lower in this study than has been reported previously.\(^{(3,4)}\) Rural families in this study owned combination type firearms to the same degree as families in nonrural areas, but children in rural areas were more likely to be exposed to rifles and shot guns, whereas children in nonrural areas were more likely to be exposed to handguns. One study found that handguns were the most common weapon used among children and adolescents who sustained serious firearm-related injuries and that the rate of injuries were 10-fold higher in urban areas than in rural areas.\(^{(20)}\) Our study findings indicate that although approximately one third of both rural and nonrural firearm owners store their firearms safely, rural families are more likely to own firearms than families in nonrural areas. These findings, taken in the context of previous studies,\(^{(21)}\) suggest that interventions for all firearm owners to store firearms safely are needed, regardless of where they live.

The risk to children for a firearm-related injury is associated with firearm storage.\(^{(19-22)}\) Previous studies have recommended (1) storage in a locked cabinet/gun safe or with a gun lock, (2) unloaded, (3) with bullets and shells stored in a separate location, and (4) with ammunition locked.\(^{(8)}\) In this study, we examined 3 of these storage patterns and noted that they varied by the types of firearms owned. Families who owned only rifles and shotguns were more likely to store guns in a hidden place other than a locked cabinet but to store ammunition separate from their guns. Handgun-only owners were more likely to store guns locked or use a gun lock but least likely to store bullets separately from their handgun. Families who owned both handguns and long guns were the least likely to use any locking procedures with guns in the home. Although there were significant differences in types of firearms owned, rural versus nonrural residence did not seem to influence safe storage practices. Moreover, our findings suggest that firearm family socialization resulted in parents’ having a less cautious approach to their family’s storage patterns. The multivariate analysis suggests that this was an important factor, in contrast to rural or nonrural region, in identifying parents who were at higher risk of unsafe firearm storage. Also, families who had younger children and owned only long guns were much more likely to store guns safely than any other subgroup. It seems that families may consider that using only 1 safe storage technique is sufficient and that the firearm type(s) owned affects which of these storage techniques is used. Although pediatricians should encourage parents to use all 3 safe storage practices in their homes, parents who are reluctant to comply with these recommendations should be encouraged at least to lock all of their firearms in a locked cabinet or gun safe and/or use gun locks such as cable locks.\(^{(17)}\)

Limitations of this study include that firearm patterns are self-reported. A previous study revealed that individuals who are not the firearm owner underestimate the unsafe storage practice and even the presence of firearms.\(^{(13)}\) To increase validity of reported safe storage, we included in regression analysis only respondents who reported being the firearm storage decision-maker. We noted that both of these groups (those who reported firearm presence in the home and the subset of firearm storage decision-makers) did not differ in their reports of storage patterns. Our overall gun ownership rate was lower than is represented by other national studies, which might have been because of high minority representation in the study.\(^{(16)}\) In addition, social desirability could have influenced who participated in the study. Last, we defined rural and nonrural status on the basis of location of the pediatric practice, which might have resulted in some rural families’ being classified as nonrural; however, it is unlikely that the opposite misclassification occurred (suburban or urban families traveling to a rural-based provider).

**CONCLUSIONS**

If guns will not be removed from homes where children live and play, then the safe storage of those guns becomes a health priority for the well-being of children. Primary care providers need to understand better not only whether firearms are in the home but also which types are present. This should inform a tailored safe storage counseling approach for gun-owning families who are at increased risk for not using safe storage practices.

**ACKNOWLEDGMENTS**

This study was supported by National Institute of Child Health and Human Development grant HD 42260, the Agency for Healthcare Research and Quality, the Robert Wood Johnson Generalist Faculty Scholars Program, the AAP’s Friends of Children Fund, and the Wachovia Foundation. In-kind support was provided through the US Department of Justice.

We especially appreciate the efforts of the PROS practices and practitioners. The pediatric practices or individual practitioners who enrolled participants in this study are listed here by AAP chapter: Alaska: Anchorage Pediatric Group, LLC (Anchorage) and Joy Neyhart, MD (Juneau); Alabama: Pediatric Care Group (Montgomery); Arizona: Orange Grove Pediatrics (Tucson) and Tanque Verde Pediatrics (Tucson); California-1: Arthur S. Dover, MD (Freedom), Palo Alto Medical Foundation (Palo Alto), and Pediatric & Adolescent Medical Associates (Salinas); Colorado: Rocky Mountain Health Centers, North (Denver) and Lamar Pediatrics (Lamar); Con-
REFERENCES


14. Connor SM. The association between presence of children in
the home and firearm-ownership and -storage practices. *Pediatrics*. 2005;115(1). Available at: www.pediatrics.org/cgi/content/full/115/1/e38


Chronic Ventilator Need in the Community: A 2005 Pediatric Census of Massachusetts

Robert J. Graham, MDa, Eric W. Fleegler, MD, MPHb, Walter M. Robinson, MD, MPHc

aDivision of Critical Care, Department of Anesthesiology, Perioperative, and Pain Medicine, and bDivision of Emergency Medicine, Department of Medicine, Children’s Hospital Boston, Boston, Massachusetts; cDivision of Respiratory Medicine, Department of Pediatrics, IWK Health Centre, Dalhousie Medical School, Halifax, Nova Scotia, Canada

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

ABSTRACT

OBJECTIVES. The purpose of this study was to describe the population of children with chronic mechanical ventilation in Massachusetts and their patterns of medical care.

PATIENTS AND METHODS. Investigators surveyed all of the Massachusetts home ventilator clinics, pediatric pulmonary services, hospital-based pediatric services for special health care needs, insurers, home care vendors, nursing agencies, the Massachusetts Department of Public Health, selected individual providers, and rehabilitation and long-term care facilities providing services to children with chronic respiratory support needs. Support was defined as daily use of noninvasive, negative-pressure, or invasive/transtracheal ventilators. Subsequent matching of demographic data, including date of birth, zip code, and gender supported maximal census yield without duplications. Geographic information systems were used to create distribution maps and estimate distances between children with chronic mechanical ventilator needs and key resources.

RESULTS. A total of 197 children were identified as requiring chronic mechanical respiratory support in Massachusetts in 2005, which was a nearly threefold increase in this population in the 15-year interval since the last census. Congenital or perinatal-acquired neurologic or neuromuscular disorders constituted the majority of primary diagnoses (n = 107 [54%]). Chronic lung disease attributed to prematurity represented only 7% of the sample.

CONCLUSIONS. Children receiving chronic mechanical respiratory support are a growing population. The shift in underlying diagnoses from pulmonary disease to neurogenic respiratory insufficiency has implications for hospital and community-based providers from all disciplines in extending services to the home setting. Barriers encountered when performing this study, however, reflect an overall lack of coordination among the many individuals and agencies involved in their care. Coordinated and centralized care efforts require a clear and managed flow of information; census reports such as this one are only the beginning. Direct needs
CHILDREN WITH SPECIAL health care needs (CSHCN) represent an increasing segment of the population.1,2 Advances in medical therapies and expansion of the supportive care options have contributed to the survival of a subpopulation of children with highly complex medical conditions.3 Increasing the levels of support from advanced technologies, however, is not without burden or risk. CSHCN have an increased risk of unscheduled PICU admission.4 Dosa et al4 reported that those with previous “technology-assisted care” had a relative risk of 375 for PICU admission versus children without chronic illness. Slonim et al5 found that CSHCN and technology dependence placed children at significantly higher risk of reported medical errors during hospitalization. For children whose illness requires chronic use of mechanical ventilators, the struggle to provide safe, quality care in the least restrictive environment challenges caregivers from all disciplines.6

In 1980, Burr et al7 documented 14 children with ventilator dependence in Massachusetts. Subsequent censuses in 1987 and 1990 by Palfrey and colleagues8,9 through “Project School Care” documented significant increases with ~220 children in Massachusetts with tracheostomy in place; between 70 and 77 children were estimated to be supported by chronic mechanical ventilation. At the time, these investigators projected that advances in the use of pulmonary surfactant in the neonatal period, as well as increasing sophistication in early neonatal care, would reduce the number of prospective children requiring chronic ventilator support. Such a projection assumed that prematurity would remain the primary insult leading to the need for chronic ventilation.

There are now a host of childhood conditions that can lead to the final common pathway of ventilator dependence. Population studies in other countries10–13 have shown that children with neuromuscular diseases, central hypventilation syndromes, spinal cord injuries, and craniofacial anomalies now compose the majority of children with tracheostomy and long-term ventilation needs. In addition to the shifting etiologies of chronic respiratory failure, advances in ventilator technology (eg, portable vents and better flow-trigger mechanisms) and extension of long-term respiratory support to other populations14,15 are likely to have changed the composition of the group first observed by Burr et al7 then canvassed by Palfrey.5,9 A new census would help care systems and providers respond to current needs of this growing population.

Given the shifting landscape of clinical care for this increasingly complex population of children, we hypothesized that the number of children requiring long-term support is greater than in previous surveys, the underlying disorders are no longer principally chronic lung disease, and the majority of patients are in the home care setting. Efforts to improve the care, safety, and quality of life for children with technology dependence and their families are contingent on a better understanding of the population. The purpose of this study was to describe the population of children with chronic mechanical ventilation needs in Massachusetts and the patterns of medical care for these children.

METHODS

Study Population and Methods

Mechanical support needs were predefined as follows: (1) daily use of noninvasive ventilator support, including facial continuous or biphasic positive airway pressure (CPAP or BiPAP); (2) negative-pressure ventilation; or (3) chronic invasive/transtracheal ventilators. We will use the term “transtracheal” to refer to ventilation delivered by tracheostomy. The amount of time on ventilator support was set at a minimum of 6 hours per day.

We acknowledge that any amount of mechanical support implies a degree of medical fragility and the need for vigilant care and monitoring and is accompanied by technical equipment burdens.

Children whose chronic respiratory support needs were limited to oxygen supplementation alone were not included in the study. We recognize that the pulmonary and cardiac conditions for these children also impact their lives and development. Greater degrees of technology, however, may serve as surrogates for disability severity,16 suggest the need for additional therapeutic services, and present logistic challenges for community providers because of physical and safety constraints. Individuals past their 22nd birthday were also excluded, because this defines the principal end point of childhood educational services and the transition from Medicaid to Medicare reimbursement programs.

Investigators contacted all of the Massachusetts home care vendors, insurers, home nursing agencies, and rehabilitation and long-term care facilities that provide services to children with chronic respiratory support needs. An additional list of physicians and other providers was generated from screening of all home ventilator clinics, pediatric pulmonary services, hospital-based pediatric services for CSHCN, and critical care providers, as well as through inquiries at the Massachusetts Consortium for Children with Special Healthcare Needs. The Massachusetts Department of Public Health was contacted as the oversight body for the state Early Intervention Program, which would enroll any child ≤3 years old. Sampling from multiple sources and subsequent matching of demographic data, including date of birth, zip code, and gender, supported maximal census yield.
without duplications, as well as identification of potential overlap in services (Fig 1).

**Study Outcome and Data Analysis**

The primary end points were an accurate assessment of the numbers of children throughout Massachusetts who require chronic, mechanical respiratory support. Secondary end points were characterization of the basic needs of this population and the extent of statewide resource use through evaluation of basic demographics (e.g., age, gender, and zip code), care site, hospitalization history within 12 months, nature of underlying respiratory insufficiency, duration of support needs, and description of actual respiratory support systems (e.g., ventilators, monitoring, and identified medical providers).

Data were entered into an SPSS 14.0 (SPSS, Inc, Chicago, IL) database for recording and analysis. Random sampling methods to equal 10% of all of the database entries were instituted to test for consistency of data entry. Primary and secondary diagnoses were recorded. The principal condition leading to respiratory insufficiency was categorized as congenital or inherited neuromuscular disorder, spinal cord injury, chronic lung disease related to prematurity, other chronic lung disease, craniofacial or upper airway abnormality, and other. Age was calculated on the basis of date of birth and date of response by the survey participant.

Descriptive statistics were computed for the entire sample and for subgroups. \( \chi^2 \) tests were used to compare the support level groups with respect to categorical variables, such as gender and care site. Analysis of variance was used to compare the ages of the patients in each support level. Geographic information systems (ArcGIS 9.1, ESRI, Inc, Redlands, CA) mapped the locations of children with chronic mechanical ventilator needs and key resources by the latitude and longitude of the centroid of their zip code. Microsoft Excel 2003 (Microsoft, Redmond, WA) was used to calculate the distances between patients and between patients and tertiary pediatric facilities using a great circle distance calculation, which accounts for the curvature of the earth.17

An a posteriori analysis of child characteristics was performed to identify those associated with need for an acute care admission in the 6-month period before census. Fisher’s exact tests were used to assess whether admission to an acute care hospital was associated with mechanical support modality, primary provider type, number of different monitoring methods, diagnostic category, age (3 levels: <4, 4–13, and ≥14 years), or gender. A multivariate logistic regression was also used to look at the association between the need for acute care hospitalization and these factors simultaneously, adjusting for age and gender.

The Children’s Hospital Boston Committee on Clinical Investigation reviewed and approved the study protocol. Because this was an observational study without an intervention and without the presentation of any identifying data, the committee ruled that informed consent was not necessary. Because of the small size and nature of the population of interest, there was the possibility of identification of individuals even through limited identifiers. Participants entered into a “limited data set use agreement” with the investigators in accordance with the Health Insurance Portability and Accountability Act. Each covered entity also reviewed the study internally, requesting supplemental study material as needed. Exchange of data with the Massachusetts Department of Public Health regarding Early Intervention Program enrollees was limited to assure compliance with Family Educational Rights and Privacy Act regulations. Survey responders were not asked for additional information, and the investigators did not have access to clients. Individual children or families were not contacted at any time. All of the data from this descriptive study were

---

**FIGURE 1**

Survey method. a Category included durable medical equipment companies/home care vendors, home nursing agencies, and the Massachusetts Early Intervention Program. b Sets of twins were identified by survey sources.
pooled and reported anonymously to assure protection of confidentiality.

RESULTS
A total of 197 children were identified as requiring chronic mechanical respiratory support in Massachusetts in 2005. Ninety-eight children (50%) received CPAP or BiPAP, whereas 97 (49%) were on invasive/transtracheal ventilators and 2 (1%) had negative-pressure devices.

Patient Characteristics
Demographic characteristics are displayed in Table 1. The median and average ages were 8.9 and 9.6 years (SD ± 6.6), respectively. Age distribution is displayed in Fig 2. There was no difference in age between children receiving transtracheal support compared with those receiving noninvasive support. Reported data on age at onset of mechanical respiratory support was limited but was skewed toward infancy and early toddlerhood (Fig 3). Gender was skewed toward boys (110 boys vs 87 girls) but was not statistically significant ($\chi^2 = 2.685; P > .10$).

A total of 138 (70%) children were identified as being cared for at home. Only 5 (3%) of the children were living in long-term care facilities. Forty-two children (21%) had been hospitalized in an acute care facility in the previous 6 months (Table 2).

Table 3 shows the nature and frequency of underlying diagnoses of the study population. Congenital, inherited, or perinatal-acquired neurologic or neuromuscular disorders constituted the majority of primary diagnoses ($n = 107$ [54%]). Chronic lung disease attributed to prematurity represented only 7% of the sample. One of the 2 patients receiving negative-pressure ventilation had chronic lung disease of prematurity, whereas the other had severe neurologic impairment after an unspecified, childhood anoxic brain injury.

Thirteen children had spinal muscular atrophy type 1, composing 7% of the sample. The average age at onset for support in this cohort of children was 4.67 months, with a range of <1 to 9 months. Eight of the children were receiving transtracheal support and 5 were receiving noninvasive CPAP or BiPAP. Of this cohort, 10 lived at home, whereas the 3 adolescents with spinal muscular atrophy were identified as living in a group home or hospital school.

Medical Resources and Management
Pediatric pulmonologists were identified as the primary providers for ventilator management in 61% of children (Table 4). General pediatricians were also reported as participating in ventilator management, but medical providers could not be identified in nearly one fourth of the cases. Monitoring modalities for all of the children are displayed in Table 5.

Figure 4 shows the geographic distribution of children throughout Massachusetts and their proximity to one another on the basis of the zip code of the primary place of residence available for 157 (78%) of 197 children. Geographic information system calculations estimate that the average distance between children with chronic mechanical respiratory support needs is 3.5 miles. Nine children have no other comparable children and an additional 6 have only 1 other comparable child living within a 10-mile radius, as demonstrated by the 10-mile buffer zones in Fig 4. The average distance to an acute-care hospital with pediatric intensive care resources is 16.7 miles, and 85.9% of identified patients live within 30 miles of such a facility.

TABLE 1  Patient Characteristics

<table>
<thead>
<tr>
<th>Characteristic Measure</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of subjects</td>
<td>197</td>
</tr>
<tr>
<td>Gender, male/female, $n$</td>
<td>110/87a</td>
</tr>
<tr>
<td>Age ($n = 196$ of 197), mean (SD), median, y</td>
<td>9.6 (±6.8), 8.9</td>
</tr>
<tr>
<td>No. (%) with transtracheostomy ventilator support</td>
<td>97 (49)</td>
</tr>
<tr>
<td>No. (%) with noninvasive CPAP or BiPAP support</td>
<td>98 (50)</td>
</tr>
<tr>
<td>No. (%) with negative pressure device</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Age of transtracheal subgroup ($N = 96$ of 97), mean (SD), median, y</td>
<td>8.8 (±7.1), 6.0b</td>
</tr>
<tr>
<td>Age of noninvasive subgroup ($N = 98$), mean (SD), median, y</td>
<td>10.3 (±6.6), 10.4b</td>
</tr>
<tr>
<td>Age at onset of mechanical ventilation ≥6 h/d ($N = 89$ of 197), mean (SD), median, y</td>
<td>6.7 (±6.1), 5.8</td>
</tr>
</tbody>
</table>

* $\chi^2 = 2.685; P > .10$.

† Student's $t$ test $= -1.523; P = .065$.

DISCUSSION
In Massachusetts, 197 children required chronic mechanical respiratory support in 2005. The underlying
cause for respiratory insufficiency was principally congenital, inherited, or perinatal-acquired neurologic or neuromuscular disorders for children receiving either noninvasive or transtracheal ventilation.

Our findings demonstrate a nearly threefold increase in this population in the 15-year interval since the study by Palfrey et al.\textsuperscript{8,9} The 1987 and 1990 censuses found 41 and \~{}70 children requiring “respirators,” respectively. In addition, the clinical characteristics of the population seem to have shifted. A higher percentage of children receive ventilation for reasons related to neurologic or neuromuscular diseases as compared with children suffering from the sequelae of prematurity. This shift may be a result of advances in neonatal care, as suggested by Palfrey et al.,\textsuperscript{8,9} but may also reflect the expansion of service to children with primary neurologic or neuromuscular impairments. Similar patterns have been documented in the United Kingdom,\textsuperscript{10} Japan,\textsuperscript{18} and Canada.\textsuperscript{13} CPAP and BiPAP usage has also expanded, which may be a product of increasing availability of equipment sized for younger children, as well as improved home care monitoring technology and changing practice patterns, which avoid tracheostomy.

The findings of the present study are important given the increasing emphasis on community inclusion and implementation of the medical home model. In this survey, 70% of children requiring chronic mechanical ventilation lived at home. Managing chronic ventilation for a child at home requires a tour de force of family and community resources, placing significant strain on fam-
ily dynamics and finances. Although many of these families will require support in the form of structural home adaptation, equipment delivery and maintenance, and transportation adaptations, these tangible needs may be less daunting than those requiring human capital. The greatest challenges for many of these families may relate to the allocation of human resources both inside and outside the family. In this context, it is worth noting that many census contributors could not identify a medical provider responsible for ventilator management, suggesting that access to appropriate physician services is not universal. Without physician supervision, coordination of care may become more difficult. When identified, bivariate analysis suggested that involvement of pediatric subspecialists was associated with less reporting of admission to an acute care facility, but this was not born out in the multivariate regression analysis. Qualitative interviews with Early Intervention Program staff revealed that physicians were not a consistent resource for the nonmedical community providers (R.J.G., D. M. Pemstein, MSW, and J. S. Palfrey, MD, unpublished data, 2007). Although allocation and availability of home care nursing or personal care assistants were not assessed as part of this study, they are also crucial factors when considering potential isolation, safety concerns, and the burdens placed on families and friends.

There are several limitations to our survey. First, we suspect that our total count is an underrepresentation of the population of children with chronic mechanical respiratory support needs. Although we received complete

### TABLE 2  Patient Care Sites

<table>
<thead>
<tr>
<th>Location</th>
<th>Current Care Site, n (%)</th>
<th>All Care Sites Within Last 6 mo, n (%)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Site at Which Patient Lived the Majority of Last 12 mo, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Home</td>
<td>138 (70)</td>
<td>146 (74)</td>
<td>136 (69)</td>
</tr>
<tr>
<td>Medical group home/school</td>
<td>12 (6)</td>
<td>17 (9)</td>
<td>12 (6)</td>
</tr>
<tr>
<td>Rehabilitation facility</td>
<td>10 (5)</td>
<td>11 (6)</td>
<td>6 (3)</td>
</tr>
<tr>
<td>Long-term care facility</td>
<td>5 (3)</td>
<td>5 (3)</td>
<td>5 (3)</td>
</tr>
<tr>
<td>Acute care facility</td>
<td>1 (1)</td>
<td>42 (21)</td>
<td>3 (2)</td>
</tr>
<tr>
<td>Not specified/unknown</td>
<td>31 (16)</td>
<td>—</td>
<td>35 (18)</td>
</tr>
<tr>
<td>Total</td>
<td>197 (100)</td>
<td>—</td>
<td>197 (100)</td>
</tr>
</tbody>
</table>

Percentages are column percentages. — indicates missing data.

<sup>a</sup> Responses do not sum to 100%, because respondents selected all of the sites in last 6 months.

### TABLE 3  Principal Diagnoses According to Mode of Support

<table>
<thead>
<tr>
<th>Principal Diagnosis</th>
<th>Transtracheal, n (%)</th>
<th>CPAP/BiPAP, n (%)</th>
<th>Negative Pressure, n (%)</th>
<th>Total, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Congenital, inherited, perinatal-acquired</td>
<td>53 (55)</td>
<td>54 (55)</td>
<td>0 (0)</td>
<td>107 (54)</td>
</tr>
<tr>
<td>neuropathic or neuromuscular disorder</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spinal cord injury</td>
<td>8 (8)</td>
<td>2 (2)</td>
<td>0 (0)</td>
<td>10 (5)</td>
</tr>
<tr>
<td>Chronic lung disease of prematurity</td>
<td>11 (11)</td>
<td>2 (2)</td>
<td>1 (50)</td>
<td>14 (7)</td>
</tr>
<tr>
<td>Chronic lung disease, other</td>
<td>8 (8)</td>
<td>4 (4)</td>
<td>0 (0)</td>
<td>12 (6)</td>
</tr>
<tr>
<td>Craniofacial or upper airway anomaly</td>
<td>8 (8)</td>
<td>21 (21)</td>
<td>0 (0)</td>
<td>29 (15)</td>
</tr>
<tr>
<td>Other</td>
<td>8 (8)</td>
<td>6 (6)</td>
<td>1 (50)</td>
<td>15 (8)</td>
</tr>
<tr>
<td>Not reported</td>
<td>1 (1)</td>
<td>9 (9)</td>
<td>0 (0)</td>
<td>10 (5)</td>
</tr>
<tr>
<td>Total</td>
<td>97 (100)</td>
<td>98 (100)</td>
<td>2 (100)</td>
<td>197 (100)</td>
</tr>
</tbody>
</table>

Percentages are column percentages.

### TABLE 4  Principal Medical Provider for Respiratory Care

<table>
<thead>
<tr>
<th>Care Provider</th>
<th>Current Ventilator Management, n (%)</th>
<th>Any Ventilator Management, Last 6 mo, n (%)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Majority Ventilator Management, Last 12 mo, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>General pediatrician</td>
<td>21 (11)</td>
<td>46 (23)</td>
<td>22 (11)</td>
</tr>
<tr>
<td>Pediatric pulmonologist</td>
<td>120 (61)</td>
<td>141 (72)</td>
<td>112 (57)</td>
</tr>
<tr>
<td>Pediatric intensivist/critical care physician</td>
<td>1 (1)</td>
<td>12 (6)</td>
<td>4 (2)</td>
</tr>
<tr>
<td>Pediatric ear/nose/throat or otolaryngologist</td>
<td>0 (0)</td>
<td>4 (2)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Neurodevelopmental pediatrician</td>
<td>2 (1)</td>
<td>5 (3)</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Other</td>
<td>9 (5)</td>
<td>13 (7)</td>
<td>10 (5)</td>
</tr>
<tr>
<td>Not specified</td>
<td>44 (22)</td>
<td>—</td>
<td>47 (24)</td>
</tr>
<tr>
<td>Total</td>
<td>197 (100)</td>
<td>—</td>
<td>197 (100)</td>
</tr>
</tbody>
</table>

Percentages are column percentages. — indicates missing data.

<sup>a</sup> Responses do not sum to 100%, because respondents selected all providers in last 6 months.
data from many sources (hospital-based providers, long-
term care and rehabilitation centers, community physi-
cians, and the Massachusetts Early Intervention Pro-
gram) and one of the third-party payors, there were
sources of information that were limited. In particular,
we were disappointed that we could not access informa-
tion from the durable medical equipment companies and
the state-based insurer MassHealth to complete this cen-
sus. Secondly, census data were collected over an
8-month period. This lengthy registry reflected time
needed to address source-specific requirement for re-
lease of data, external data retrieval, and logistics of
exchange. Attrition because of aging out or death and
additions to the populations during the period of the
census were likely but should not have impacted on
approximate prevalence estimates. Finally, because of
differences in the interpretation of limited personal iden-
tifiers between investigators and sources, the demo-
graphic data were not complete in every identified case.
Some census contributors also provided incomplete data,
including care site, age at onset of support, medical
resources, and management, because of deficiencies in
their own records. As a consequence, bivariate and mul-
tiple regression analysis of patient characteristics and
need for admission to an acute care facility were also
limited and do not reflect the individual patient who
may have required repeated or prolonged admission.
The significant relationship identified between the num-
ber of monitoring modalities and acute hospitalization
likely reflects the underlying degree of illness.
This census demonstrates that children receiving
chronic mechanical respiratory support are a small but
growing population with implications for hospital and
community-based providers from all disciplines. The shift
in underlying diagnoses from pulmonary disease to neu-
rogenic respiratory insufficiency suggests a change in the
population needing expert services to remain in the home
setting. Complex care requires communication and coor-
dination; for these children and their families, communi-
cation and collaboration between multiple medical sub-
specialties, therapists, and home/community providers is
essential. Yet, the difficulty of performing a simple census
of these children suggests an overall lack of coordination
among the many individuals and agencies involved in their
care. Difficulty in accessing even limited clinical status
records, as we encountered in performing this census, will
hamper both recognition of the need for coordinated pro-
grams for this population and the implementation of such
programs. It also bears recognition that these barriers to
health services research were encountered in a region

<table>
<thead>
<tr>
<th>Variable</th>
<th>Transtracheal, n (%)</th>
<th>Noninvasive Ventilation, n (%)</th>
<th>Negative Pressure Ventilation, n (%)</th>
<th>Total, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulse oximetry (saturation)</td>
<td>73 (75)</td>
<td>55 (56)</td>
<td>2 (100)</td>
<td>129 (66)</td>
</tr>
<tr>
<td>End tidal CO₂</td>
<td>5 (5)</td>
<td>2 (2)</td>
<td>0</td>
<td>7 (4)</td>
</tr>
<tr>
<td>Cardiovascular/respiratory leads</td>
<td>11 (11)</td>
<td>1 (1)</td>
<td>0</td>
<td>12 (6)</td>
</tr>
<tr>
<td>Continuous direct observation</td>
<td>20 (21)</td>
<td>6 (6)</td>
<td>1 (50)</td>
<td>27 (14)</td>
</tr>
<tr>
<td>None</td>
<td>0</td>
<td>17 (17)</td>
<td>0</td>
<td>17 (9)</td>
</tr>
<tr>
<td>Other, not specified</td>
<td>3 (3)</td>
<td>3 (3)</td>
<td>0</td>
<td>6 (3)</td>
</tr>
</tbody>
</table>

Percentages are column percentages. Responses do not sum to 100%, because respondents selected all that apply.

FIGURE 4
Geographic distribution of children with chronic me-
chanical respiratory support. This map is projected in the
State Plane Coordinate system: Massachusetts Mainland
based on the North American Datum of 1983, which
minimizes distortion of the state but does project the
10-mile buffers, which are true circles, as ellipses. * So-
cially isolated is defined as 0 to 1 child within a 10-mile
radius.
where there is an abundance of academic institutions and tertiary care pediatric facilities. It may be that electronic medical charts will improve the situation, but we are not optimistic. The difficulty of counting these children is more discouraging in light of the evidence that innovative and coordinated programs of care for medically vulnerable children are successful. Comprehensive programs, such as the Ventilator Assisted Children’s Home Program in Pennsylvania, have established standards of care, demonstrated cost savings, and decreased mortality among enrollees.19–21

Coordinated and centralized care efforts require a clear and managed flow of information; census reports such as this one are only the beginning. Consistent identification of these children is pivotal when providing services. Direct needs assessments and quality-of-life surveys from families are needed to design and implement programmatic changes and advocacy efforts. Delination of shortages of homecare nursing is necessary as are innovative programs and policy to foster growth in this area. Use of medical passports,22 improved discharge planning,23 coordinated follow-up,19,24 support for homecare nursing, creative use of Web-based technology, parent-to-parent support groups, and other strategies should be studied to enhance patient care, outcomes, and satisfaction and to limit the burden of care.

ACKNOWLEDGMENTS

This study was supported by a grant from the Thoracic Foundation (Boston, MA).

We thank Peter Forbes, MA (Clinical Research Program, Children’s Hospital Boston) for statistical assistance and Ingrid Liff, BA, Judith S. Palfrey, MD, and all of the survey participants for contributions to the implementation and completion of this study. The Massachusetts Consortium for Children With Special Health Care Needs also facilitated our contacts with numerous providers and insurers.

REFERENCES

11. Fraser J, Mok Q, Tasker R. Survey of occupancy of paediatric intensive care units by children who are dependent on ventilators. BMJ. 1997;315:347–348
To Clean or Not to Clean: Effect on Contamination Rates in Midstream Urine Collections in Toilet-Trained Children

Suzanne Vaillancourt, MDa, David McGillivray, MDa, Xun Zhang, PhDb, Michael S. Kramer, MDb

aDivision of Pediatric Emergency Medicine and bDepartments of Pediatrics and Epidemiology and Biostatistics, Montreal Children’s Hospital, and cMontreal Children’s Hospital Research Institute, McGill University, Montreal, Quebec, Canada

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

ABSTRACT

OBJECTIVE. Urinary tract infection is one of the most common bacterial infections among children. Difficulty in specimen collection and interpretation of inadequately collected specimens may contribute to misdiagnosis of urinary tract infection. Our objective was to assess the effect of perineal/genital cleaning on bacterial contamination rates of midstream urine collections in toilet-trained children.

METHODS. We conducted a randomized trial in toilet-trained children who presented to a tertiary care pediatric emergency department between November 1, 2004, and October 1, 2005. All toilet-trained children who were between the ages of 2 and 18 years and had a midstream urine sample requested were eligible. Those whose parents consented were cluster-randomized by week to either cleaning or not cleaning the perineum with soap. The risk for a contaminated urine culture (defined as growth of $<10^8$ colony-forming units per liter [$<10^5$ colony-forming units per milliliter] of a single organism or a mix of $\geq 2$ organisms) and the risk for a positive urinalysis (defined as a positive leukocyte esterase and/or nitrites on dipstick or $\geq 5$ white blood cells per high-powered field on a standard microscopic examination) were analyzed by intention to treat.

RESULTS. A total of 350 children were enrolled. The rate of contamination in the cleaning group was 14 (7.8%) of 179 vs 41 (23.9%) of 171 in the noncleaning group. Children who were randomly assigned to cleaning were less likely to have a positive urinalysis (37 of 179 [20.6%]) than those in the noncleaning group (63 of 171 [36.8%]).

CONCLUSIONS. Urine contamination rates are higher in midstream urine that is collected from toilet-trained children when obtained without perineal/genital cleaning. Cleaning may reduce the risk for returning for repeat cultures and for receiving unnecessary antibiotic treatment and investigations.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2392
doi:10.1542/peds.2006-2392

Key Words
urinary tract infection, midstream clean catch, contamination rates

Abbreviations
UTI—urinary tract infection
ED—emergency department
CFU—colony-forming unit(s)
RR—relative risk
OR—odds ratio
CI—confidence interval
GLIMMIX—generalized linear matrix model

Accepted for publication Nov 21, 2006
Address correspondence to Suzanne Vaillancourt, MD, Montreal Children’s Hospital, Place Tourelles, 2222 Blvd Rene Levesque Ouest, First Floor, Room 1311-C, Montreal, Quebec, Canada H3H 1P3. E-mail: suzanne.vaillancourt@mail.mcgill.ca

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
Urinary tract infection (UTI) is one of the most common bacterial infections among children. UTIs are an important cause of chronic morbidity in children; long-term complications include hypertension and reduced renal function secondary to renal scarring. The accurate diagnosis of a UTI is necessary to ensure appropriate therapy for infected children and to avoid unnecessary therapy and prevent hospital admission for additional evaluation in noninfected children. A urine culture result is considered the gold standard for diagnosis of a UTI, but difficulty in specimen collection and interpretation of inadequately collected specimens may contribute to its misdiagnosis in children.

Few studies have addressed the method of collecting a midstream urine specimen in outpatient children. The “clean-catch” midstream void technique has been the method of choice in adults since the late 1950s. The clean-catch midstream technique, however, is time-consuming to explain, frequently performed incorrectly, and associated with increased costs. Several studies in adults have reported no difference in contamination rates between midstream clean-catch and midstream non–clean-catch urine samples. Authors of these studies hypothesized that cleaning does not lower contamination rates because the urethra is irritated during cleaning. In toilet-trained children, only 3 studies have evaluated the effect of cleaning on bacterial contamination rates in midstream specimens. All 3 studies showed no difference in contamination rates, but none was randomized; therefore, their results may be confounded by differences in those cleaned versus uncleaned. In this article we describe the first randomized trial, to our knowledge, to address this question.

METHODS

Study Design and Patients

We conducted a randomized trial in toilet-trained children who presented to the emergency department (ED) of a tertiary care pediatric center between November 1, 2004, and October 1, 2005. All toilet-trained children who were between the ages of 2 and 18 years and had a midstream urine sample requested by a physician or triage nurse for any medical reason were eligible for the study. Children were excluded when they were not toilet-trained or their parents believed that they would be unable to comply with the collection technique because of developmental delay. Informed consent was obtained by either the research assistant who worked in the ED or the physician who requested the urine sample. The hospital research ethics board approved the study.

At the beginning of each week throughout the study period, the method for collecting a urine specimen for all children who presented to the ED during that week was randomly assigned by the study investigators. A card stating either “cleaning” or “noncleaning” was randomly chosen from a hat by the study investigator on each Monday morning at 8 AM. Multiple posters were then displayed to inform the ED staff as to whether it was a cleaning or noncleaning week. In addition, during the cleaning week, the soap was given to the children and their parents, and during a noncleaning week, the soap was not given to the children and their parents. Allocation to either the cleaning group or the noncleaning group was the same for all children who presented during that week. The nurses and physicians were not blinded to the collection technique. After informed consent was obtained, verbal and written instructions for obtaining the urine specimen were given to the child and parent(s).

During the cleaning week, the child was given liquid soap (Isoderm; Lalema, Montreal, Quebec, Canada), several gauze pads, a sterile urine-collection container, and an instruction sheet. The child and/or the parent was instructed to spread the labia (for girls) or retract the foreskin (for noncircumcised boys), to clean the urethral meatus and perineum with gauze and liquid soap twice (wiping from front to rear in girls), to urinate into the toilet, and, midway through urination, to collect the urine into the sterile container. During the noncleaning week, the child was given a sterile urine-collection container and an instruction sheet. The child and/or the parent was instructed to spread the labia (for girls) or retract the foreskin (for noncircumcised boys), to urinate into the toilet, and, midway through urination, to collect the urine into the sterile container.

A questionnaire was administered with the urine-collection instructions to the parent of each participating child to document the age, gender, circumcision status (for boys), antibiotic use in the previous 2 weeks, and previous renal problems (other than a previous UTI). Recent antibiotic use was documented because bacterial colony counts may fall below the characteristic range for an infection when an antimicrobial agent is present in the urine. To assess compliance with the collection instructions, we also asked whether the child cleaned with soap before collecting the urine sample. We did not collect data on the presence or absence of urinary symptoms. Information that was recorded from the urinalysis included the presence of nitrites or leukocyte esterase in fresh urine and the presence or absence of bacteria and the number of white blood cells per high-powered field on microscopy of spun urine.

Laboratory Methods and Definitions

On the basis of previous studies, the standard of practice at our institution is to perform the urine dipstick on all urine specimens. When the dipstick indicates a positive finding for any component (including leukocyte esterase, nitrites, ketones, protein, glucose, or blood), a microscopic urinalysis is performed. A positive urinalysis is defined as a positive leukocyte esterase and/or nitrites.
on dipstick or ≥5 white blood cells per high-powered field on a standard microscopic examination done after centrifuging the urine. All samples with a positive urinalysis were sent for culture. Given that false-negative urinalysis results in toilet-trained children are rare, urine samples are not sent for culture when the urinalysis is negative.

Urine specimens were sent to the microbiology laboratory in sterile containers. Standard quantitative culture was performed by laboratory technologists. Urine was plated onto MacConkey agar plates and sheep-blood agar plates with a 0.01-mL calibrated loop. Cultures were incubated for 48 hours at 35°C and examined daily for growth. Microbiologists at the laboratory were blinded to the randomization week and hence to the cleaning or noncleaning status of each child’s specimen. A positive urine culture was defined as growth of a single pathogenic organism with \(10^8\) colony-forming units (CFU)/L (\(10^5\) CFU/mL) of urine. Urinary tract pathogens included *Escherichia coli*, *Proteus*, *Klebsiella*, *Enterococcus*, *Enterobacter*, and *Pseudomonas*. Nonpathogenic organisms included *Lactobacillus*, *Corynebacterium*, and coagulase-negative staphylococci. A contaminated urine culture was defined as a culture growing a single organism with \(<10^8\) CFU/L (\(<10^5\) CFU/mL) or a mix of 2 or more organisms. A negative culture was defined as no growth. For assessment of whether the contamination rate was affected by the time of day when the sample was plated in the microbiology laboratory, the time of collection of the urine sample was recorded.

**Statistical Analysis**

We estimated the number of children needed to detect a clinically important difference of 20%, based on an \(\alpha\) value of .05 and a power of 80%, as 182 in each group. Relative risks (RR) and their 95% confidence interval (CI) for contamination and for a positive urinalysis were calculated using SPSS 11.0 (SPSS, Chicago, IL). Analysis was based on intention to treat. To test for a difference in the culture result between the cleaning and noncleaning groups among samples with a positive urinalysis, we calculated a \(\chi^2\) test statistic. Because the intervention was randomized by week rather than by individual patient, a generalized linear mixed model (GLIMMIX) was also used to account for potential within-week clustering of outcome results and to adjust for potential baseline imbalances in gender, age, and history of renal problems. The interaction terms for treatment by gender and age were also included in the GLIMMIX model.

**RESULTS**

During the study period, 350 children who met the eligibility criteria were enrolled (Fig 1). The majority were girls (211 [60%]). Sixty-five percent of the boys enrolled were noncircumcised. Preschool children (aged 2–5) composed 45% of the study sample. Table 1 displays the baseline characteristics of all participating children and shows that the 2 intervention groups were similar in these characteristics. Reported adherence to the urine-collection instructions was excellent: >90% of parents stated that their child followed the instructions given. The parents of 10 children in the noncleaning group indicated on the questionnaire that they cleaned their child with soap before giving the urine sample, and 15 parents of children in the cleaning group stated that they did not clean with soap despite being instructed to do so.

The overall prevalence of UTI in the study population was 7%, with the most common pathogen, *Escherichia coli*, growing in 19 (83%) of 23 of the positive cultures. Two other pathogens isolated were *Enterococcus* (\(n = 1\)) and *Proteus* (\(n = 3\)) species. Ninety-six percent of the positive cultures were seen in girls, with a mean age of 8 years. The majority of contaminated specimens grew a mix of 2 or more species (\(n = 34\)). The remaining contaminated cultures grew either 1 organism \(<10^7\) CFU/L.
(<10^4 CFU/mL; n = 15) or nonpathogenic organisms, including Corynebacterium (n = 3) and coagulase-negative Staphylococcus epidermidis (n = 3). Of the 55 contaminated urine specimens, 21 were collected during the day and 34 were collected during the evening. None of the specimens that were collected overnight was contaminated. The contamination rate in the cleaning group was 14 (7.8%) of 179 compared with 41 (23.9%) of 171 in the noncleaning group (RR: 0.37; 95% CI: 0.19–0.57; Table 2). Urine specimens that were obtained with cleaning were less likely to have a positive urinalysis (37 of 179 [20.6%]) than those that were obtained in the noncleaning group (63 of 171 [36.8%]) (RR for positive urinalysis: 0.56; 95% CI: 0.40–0.79; Table 2). In the GLIMMIX model used to account for potential clustering within weeks, the effect was similar; the OR for contamination in the cleaning group was 0.27 (95% CI: 0.14–0.51), and the OR for a positive urinalysis in the cleaning group was 0.45 (95% CI: 0.28–0.72). In the cleaning group, the predictive value of a positive urinalysis for UTI was 15 (40.5%) of 37, compared with 8 (12.7%) of 63 in the noncleaning group (P = .001). Sensitivity, specificity, and negative predictive value could not be calculated, because only specimens with positive urinalyses were sent for culture. When a positive urinalysis culture was defined as ≥10^7 CFU/L (≥10^5 CFU/mL) instead of ≥10^8 CFU/L (≥10^5 CFU/mL), only 1 urine sample that previously was classified as contaminated would be reclassified as positive and the contamination rates would be almost identical: 14 (7.8%) of 179 in the cleaning group versus 40 (23.4%) of 171 in the noncleaning group (RR: 0.38; 95% CI: 0.20–0.59).

We also assessed the treatment effect after stratification by gender, age, and circumcision status. In the age group 2 to 5 years, 15.3% in the noncleaning group were contaminated versus 10.3% in the cleaning group. In the age group 6 to 18 years, 28.6% of samples in the noncleaning group were contaminated versus 6.3% in the cleaning group. The interaction between age and treatment was statistically significant (P = .04), whereas the interaction between gender and treatment was not (P = .15). Only 4 boys had contaminated urine specimens, 3 of whom were uncircumcised (1 in the cleaning group and 2 in the noncleaning group), with no circumcised boys in the cleaning group, thereby preventing testing for interaction by circumcision status.

**DISCUSSION**

In this cluster-randomized trial, we observed that children who were randomly assigned to receive genital/perineal cleaning with soap before urine collection had lower rates of contaminated urine specimens and of positive urinalyses. Moreover, the predictive value of a positive urinalysis for UTI was significantly higher in the cleaning group. We are aware of only 3 previous studies that compared methods of obtaining a midstream urine sample in children, none of which found cleaning to reduce the contamination rate.14–16 The definition of a contaminated urine sample used in those studies was similar to that used in our study, but none used randomized allocation, their sample sizes were smaller, and the collection techniques were not standardized.

The results of our study and other studies that assessed rates of contamination depend on the definition of a contaminated urine specimen. Since the 1950s, the traditional <10^8 CFU/L (<10^5 CFU/mL) cutoff point has been used by clinicians to denote a positive culture.20 The prevailing view is that a mixed culture in an uncomplicated outpatient likely indicates contamination. Low levels (<10^7 CFU/L [<10^4 CFU/mL]) of organisms that commonly are found on the skin and external genitalia are considered to be contaminants. After reviewing the literature,13,14,20–24 we defined a positive urine culture as growth of ≥10^8 CFU/L (≥10^5 CFU/mL) of urine from a clean-catch midstream specimen. Our overall results were very similar when we defined a positive culture as ≥10^7 CFU/L (≥10^4 CFU/mL): all but 1 of the contaminated specimens were attributable either to mixed specimens or to nonpathogenic organisms with <10^7 CFU/L (<10^4 CFU/mL). In previous studies in adults, contamination rates for clean-catch midstream urine specimens have ranged from 0% to 32%.25–27

**TABLE 1** Characteristics of Study Sample

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cleaning Group (N = 179), n (%)</th>
<th>Noncleaning Group (N = 171), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male gender</td>
<td>72 (40)</td>
<td>67 (39)</td>
</tr>
<tr>
<td>Age, y</td>
<td>81 (45)</td>
<td>76 (44)</td>
</tr>
<tr>
<td>2–5</td>
<td>70 (39)</td>
<td>59 (35)</td>
</tr>
<tr>
<td>6–12</td>
<td>28 (16)</td>
<td>36 (21)</td>
</tr>
<tr>
<td>Collection time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8–12 AM</td>
<td>173 (98)</td>
<td>164 (98)</td>
</tr>
<tr>
<td>12–8 AM</td>
<td>4 (2)</td>
<td>4 (2)</td>
</tr>
<tr>
<td>Circumcised</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>22 (31)</td>
<td>27 (40)</td>
</tr>
<tr>
<td>No</td>
<td>50 (69)</td>
<td>40 (60)</td>
</tr>
<tr>
<td>Recent antibiotic use</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>25 (14)</td>
<td>30 (18)</td>
</tr>
<tr>
<td>No</td>
<td>154 (86)</td>
<td>141 (82)</td>
</tr>
<tr>
<td>Renal problem</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>17 (9)</td>
<td>12 (7)</td>
</tr>
<tr>
<td>No</td>
<td>162 (91)</td>
<td>159 (93)</td>
</tr>
</tbody>
</table>

**TABLE 2** Risk for Contamination and a Positive Urinalysis

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Cleaning Group (N = 179)</th>
<th>Noncleaning Group (N = 171)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contaminateda</td>
<td>14</td>
<td>41</td>
<td>55</td>
</tr>
<tr>
<td>Noncontaminated</td>
<td>165</td>
<td>130</td>
<td>295</td>
</tr>
<tr>
<td>Positive urinalysisb</td>
<td>37</td>
<td>63</td>
<td>100</td>
</tr>
<tr>
<td>Negative urinalysis</td>
<td>142</td>
<td>108</td>
<td>250</td>
</tr>
</tbody>
</table>

a RR: 0.37 (95% CI: 0.19–0.57).
b RR: 0.56 (95% CI: 0.40–0.79).
In exploratory posthoc analyses, we observed a significantly larger difference in contamination rates according to treatment among children who were 6 to 18 years than among those who were 2 to 5 years. This difference in effect of cleaning may be attributable to poorer local hygiene and/or more careful cleaning in the older children.

Because urine samples with a positive urinalysis were sent for culture and because we observed more positive urinalyses among the noncleaning group, a higher percentage of samples from the noncleaning group were sent for culture; this has implications for clinical practice, because clinicians prefer to identify children with a sufficiently increased likelihood of UTI to justify presumptive antibiotic treatment while awaiting the results of the urine culture, which may not be available for 24 to 48 hours. Management decisions are therefore often based on the results of the urinalysis and can include treating the child immediately with antibiotics, waiting until the culture result is known, or obtaining a repeat collection of the urine sample.

It has been hypothesized that daytime urine (fluid diuresis) and frequent daytime voiding can result in lower colony counts. Conversely, overnight dwelling of urine in the bladder may lead to an increased concentration of bacteria. However, time of collection did not affect the contamination rates in our study.

Although written instructions were given to the child and the parent, it is difficult to verify that either cleaning or midstream collection was accurately performed. We attempted to assess the collection procedure by having the child and/or the parent complete a questionnaire to assess compliance with the collection instructions. Our procedure reflects what occurs in clinical practice, because neither the physician who requests the urine nor the nurse who gives the instructions is involved in observing or performing the collection of the urine. In any case, any unreported noncompliance with the allocated procedure should have reduced differences in contamination rates between the 2 study groups.

CONCLUSIONS

Urine contamination and positive urinalysis rates were lower in midstream urine that was collected from toilet-trained children who were randomly assigned to receive perineal/genital cleaning. We also observed a higher positive predictive value for the urinalysis in those children. Although our study was limited to children who were seen in an ED setting, our results should be generalizable to other clinical settings. It is important to make an accurate diagnosis of a UTI in children to provide appropriate treatment and evaluation for these children. However, a collection method that results in higher contamination rates is likely to lead to an increased proportion of children’s having to return for repeat cultures and/or potentially receiving unnecessary treatment and investigations. Our results therefore strongly suggest that toilet-trained children should have their perineum cleaned with soap before collecting a midstream urine specimen.

ACKNOWLEDGMENTS

This research was supported by a grant from the Montréal Children’s Hospital Research Institute.

Duncan Lejtenyi, Aileen Frew, and the nurses of the Montreal Children’s Hospital Emergency Department assisted in recruiting children into the trial. Sebastian Dube provided statistical assistance. Dr Kramer is a Senior Investigator of the Canadian Institutes of Health Research.

REFERENCES

Heelys and Street Gliders Injuries: A New Type of Pediatric Injury

Mihai Vioreanu, MRCSI, Eoin Sheehan, AFRCSI, Aaron Glynn, AFRCSI, Noelle Casidy, FRCSI, Michael Stephens, FRCSI, Damian McCormack, FRCSI

Trauma and Orthopaedic Surgery Department, Temple Street Children’s University Hospital, Dublin, Ireland

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

ABSTRACT

OBJECTIVES. Our goals were to highlight an increasing trend in orthopedic injuries in children as a result of “heeling” or “street gliding,” to describe injuries sustained by children using Heelys (HSL, Carrollton, TX) and Street Gliders (Glowgadgets Ltd, Bristol, United Kingdom), and to increase public awareness and prevent such injuries.

PATIENTS AND METHODS. We prospectively recorded the data of all roller shoes injuries referred to our department during the summer school holiday. Using a data-collection sheet, we recorded demographic data, type of injury, mechanism and place of injury, heeling or street-gliding experience, use of safety equipment, methods of treatment, and intention to continue heeling or street gliding after recovery from injury.

RESULTS. Over a 10-week period, 67 children suffered orthopedic injuries while using Heelys or Street Gliders. There were 56 girls and 11 boys with a mean age of 9.6 years. Upper limbs were the most common location of injury. Distal radius fractures were the most prevalent, followed by supracondylar fractures, elbow dislocations, and hand fractures. The majority of children suffered the injury while heeling or street gliding outdoors. Interestingly, 20% of the injuries happened while trying Heelys or Street Gliders for the first time, and 36% of the injuries occurred while learning (using 1–5 times) how to use them. None of the children used any sort of protective gear at the time of the injury. The majority of the injured children expressed their intention to continue heeling or street gliding after complete recovery from their injury.

CONCLUSIONS. Our study shows that the majority of children with injuries from heeling or street gliding are girls. We recommend close supervision of children using Heelys or Street Gliders during the steep learning curve and usage of protective gear at all times. These new types of injuries have a serious impact on child health and constitute a burden for the pediatric orthopedic service.
HEELYS (HSL, CARROLLTON, TX) are the only shoes with a removable wheel in the heel (Fig 1). The innovative shoes with wheels in the heel allow “heelers” to go from a walk to a roll by shifting body weight. They were first launched in 2000 in the United States, and since then, Heelys have become extremely popular among children all over the world. The producing company reports the distribution of >4.5 million pairs in >60 countries worldwide. The popularity of the Heelys, also called roller shoes, spawned many imitations, increasing the accessibility of children to this activity. Street Gliders (Glowgadgets Ltd, Bristol, United Kingdom) are a different version in which wheels are strapped to regular running shoes to simply combine walking with rolling only by shifting body weight (Fig 2). With the growing popularity of this new trend, an increasing number of media reports raising the awareness of a high injury risk have been released.

To change from walking to rolling while using Heelys or Street Gliders, children place 1 foot in front of the other and shift their body weight backward over the wheels in the heels. This is called, in user terms, “crocodile stance” (Fig 3). The skating position of Heelys and Street Gliders is a balanced position of a body that tends to fall backward. This is achieved by contracting the hip extensors muscles, flattening the lumbar lordosis, and pushing the shoulder and neck forward.

To further report on injury risk of this new trend, we prospectively recorded injuries resulting from “heeling” or “street gliding” that required orthopedic treatment during school summer holiday in 2006. The objective of this article is to describe the orthopedic injuries in children resulting from heeling or street gliding.

METHODS
All of the children presented to the Orthopaedic Surgery Department, Temple Street Children University Hospital, for assessment and additional treatment of fractures resulting from using Heelys and Street Gliders, were eligible for inclusion in the study. Our institution is a tertiary referral pediatric trauma unit covering a population of 130 000 children. All of the children in the area requiring orthopedic assessment or treatment are referred to our unit. This is a consecutive series of orthopedic injuries because of heeling or street gliding that we prospectively recorded during school summer holiday (July 1 to September 15, 2006). Using a data-collection sheet, we recorded demographic data (gender and age), type of injury, mechanism and place of injury, heeling or street gliding experience, use of safety equipment, methods of treatment, and intention to continue heeling or street gliding after recovery from injury. We recorded all of the data ourselves. The study was conducted in conformity with the Helsinki II Declaration, and parents of all of the children included in the study gave their consent.

Children were followed up as outpatients with regular visits to our clinic until full clinical resolution of their respective injuries. There were no patients lost to follow-up.

RESULTS
Demographic Data
Over the 10-week period, 67 children were treated for Heelys- or Street Gliders–related injuries. Ages ranged from 6 to 15 years; the mean age was 9.6 years. The majority of injured children (83.5%) were girls. During this period of time, a total of 830 children were referred to our department for assessment and treatment of traumatic fractures; thus, Heelys and Street Gliders injuries represented 8% of the workload of the orthopedic department during the study period. From the total number of injuries, 78% (50 of 67) were because of heeling, and 22% (17 of 67) were because of street gliding.
study we did not include children with minor injuries who did not require orthopedic assessment or treatment or those who were treated by the accident and emergency department only.

**Mechanism and Location of Injury**
The mechanism of injury was falling backward (55%) or forward (32%) as the user was trying to transfer the body weight and find balance on the wheels. In a few cases (13%), the injury was caused by a fall because of jumping or sudden change in direction of motion. The majority of injuries (76%) happened while heeling or street gliding outdoors (road, sidewalk, cycle lane, or playground). From the injuries that took place indoors (24%), 10 (15%) were in a shopping mall, and 7 (9%) were in the family home.

**Injury Pattern and Management**
The majority of children (58 of 67 [87%]) sustained upper-limb trauma. Distal radial fractures were the most common type of injury encountered. There were 50 distal radius fractures in 49 children (1 child sustained bilateral distal radius fractures). There were also 5 supracondylar fractures, 2 elbow dislocations, and 2 metacarpal fractures. The lower limb was also injured in 9 patients: 6 foot and ankle injuries, 2 tibia/fibula fractures, and 1 knee injury. In the presented series of injuries, there were no head injuries, and none of the injuries were life threatening. From the total number of patients referred to our department, 38% (24 of 67) required admission to the hospital for manipulation under anesthesia and cast application.

**Contributing Factors**
When asked to rate their own level of experience of heeling or street gliding, the majority (70%) of children considered themselves beginners, having used the Heelys or Street Gliders between 1 and 5 times only; 23 children (34%) were first-time users. There were 8 children (12%) with intermediate experience, and 12 children (18%) were advanced in using roller shoes. The more experienced children suffered an injury while trying to jump or suddenly change direction. None of the injured children were wearing any kind of protective gear at the time of their injury, and only 8 children (12%) were fully familiar with the user instructions provided by the manufacturer.

Interestingly, when asked about their intention of continuing heeling or street gliding after recovery from their injury, 54% (36 of 67) of the children expressed their intention to use the Heelys or Street Gliders again. This was in big disagreement with the view of their parents.

**DISCUSSION**
Here we present a consecutive series, over 10 weeks, of 67 children who suffered orthopedic injuries while using Heelys or Street Gliders. There were 56 girls and 11 boys with a mean age of 9.6 years of age in our study group. Upper limbs sustained by far the most common type of injury. Distal radius fractures were the most prevalent, followed by supracondylar fractures, elbow dislocations, and hand fractures. The majority (76%) of children suffered the injury while heeling or street gliding outdoors. Interestingly, 20% of the injuries happened while trying Heelys or Street Gliders for the first time, and 36% of the injuries were sustained while learning (using 1–5 times) how to use them. None of the children used any sort of protection gear at the time of the injury, and only 12% were familiar with the instructions of use.

Although Heelys and Street Gliders are dramatically increasing in popularity among children around the world, only 1 study reports on injuries associated with this new sport.1 The authors reported a total of 37 seri-
ous orthopedic injuries over 6 months. Their patients suffered mainly upper-limb fractures (only 1 patient had a lower-limb fracture), with high preponderance of distal radius (55%) and elbow (34%) fractures. None of their patients wore any kind of safety or protection gear while heeling. The authors noted that, unlike roller blading and cycling, Heelys provide a built-in versatility, allowing children to change from walking to rolling whenever they want to do so. Thus, children are less likely to wear protective gear at all times.

Although recreational skating can improve the health of children through exercise, participation in skating activities exposes children to a risk of injury. Multiple studies have explored pediatric injuries associated with roller skating and in-line skating.1–3 Our institution previously published 2 articles reporting on roller blade injuries.4–5 In 1998, the American Academy of Pediatrics issued a recommendation that children and adolescents wear full protective gear, including a helmet, wrist guards, knee pads, and elbow pads, during in-line skating. A decrease in injuries related to in-line skating started in 1999. The trend of decreasing in-line skating-related injuries may reflect more children wearing protective gear and/or the popularity of in-line skating decreasing.6 The National Safe Kids Campaign in the United States estimated that, in 2002, the number of children treated for roller skating–related injuries remained largely the same, at ~28,400, and the number treated for in-line skating–related injuries declined from 42,800 to 36,300.7

The Canadian Safety Council issued a consumer alert (January 2006) advising children and parents of children using Heelys to wear protective gear and to avoid heeling on roads, sidewalks, and wet surfaces. They recommended a ban on heeling in public buildings and malls, as well as school hallways and playgrounds.8 In Europe and the United States there is no official regulation or public health alert concerning the safety of using Heelys or Street Gliders.

The most common mechanism of injury is falling backward or forward. The majority of injuries occurred while children were learning how to use roller shoes. Control and balance of the body’s center of gravity is very important in wearing these roller shoes. The Canadian Safety Council, in a consumer safety alert, estimated that a month with daily practice is required to master this technique.9 At the time of the injury, none of our patients used any sort of protective gear, and only 12% were familiar with the instructions for use that were provided. The alternative mechanism of injury was falling while trying different tricks, such as jumping or suddenly changing direction while rolling. These injuries were encountered in children with advanced heeling or street-gliding skills. We noticed that a higher percentage (7 of 15 [47%]) of street gliding–related fractures required manipulation under anesthetic than did heeling-related injuries (11 of 52 [21%]). Street gliding–related fractures are higher energy injuries than heeling-related injuries. This is because of the larger diameter of the wheels in Street Gliders, allowing rolling at a higher speed than with Heelys. Most injuries happened outdoors, on streets, footpaths, or playgrounds. The majority of injured children in our study were girls. Although we do not have data on gender distribution of users of roller shoes, this could explained by a higher number of girls using Heelys or Street Gliders.

It seems that the new “walk-and-roll” generation is really enjoying the development of roller shoes. Despite their parents’ intention that their children stop using Heelys or Street Gliders, 54% of the injured children expressed their intention to continue using them after recovery of their injury.

CONCLUSIONS

We report the single largest prospective epidemiologic study of orthopedic injuries in children that resulted from the increasingly popular new activities heeling and street gliding during the school summer holiday. We highlight the need for regulatory and safety recommendations to be issued by governmental bodies and agencies regarding marketing and safe use of Heelys and similar roller shoes.

To reduce the rate of such injuries, parents buying roller shoes need to understand both the benefits and risks of this activity. Children and their parents should appreciate that injuries are particularly common in novice users and those more adventurous advanced users. Full protective gear needs to be used at all times, including a helmet, wrist guards, knee pads, and elbow pads, when using roller shoes. Wrist guards should always be worn to reduce impact forces and distribution when a child falls on the outstretched hand, because upper-limb trauma constitutes 86.5% of our cases. Special attention should be paid to the needs of novice skaters to avoid injuries. We recommend that a safe-use guide be provided with each pair of roller shoes. An induction demonstration on safe use should be provided by retail outlets at the time of purchase of these roller shoes.

REFERENCES


Comparison of the Safety and Immunogenicity of a Refrigerator-Stable Versus a Frozen Formulation of ProQuad (Measles, Mumps, Rubella, and Varicella Virus Vaccine Live)

Henry H. Bernstein, DOa, Karen Eves, BSb, Kristy Campbell, MSb, Steven B. Black, MDd, Jerry D. Twiggs, MDd, Keith S. Reisinger, MDe, Ralph M. Conti, MDd, Carl-Erik Flodmark, MDg, Lars Rombo, MDh, Stephanie Klopfer, PhDh, Florian Schödel, MDh, Jonathan Hartzel, PhDh, Barbara J. Kuter, PhD, MPHb, and the Refrigerator-Stable Formulation Study Group for ProQuad

aDartmouth Medical School, Children’s Hospital at Dartmouth, Lebanon, New Hampshire; bMerck Research Laboratories, West Point, Pennsylvania; cKaiser Permanente Vaccine Study Center, Oakland, California; dDixie Pediatrics, St George, Utah; ePrimary Physicians Research, Pittsburgh, Pennsylvania; fFoothills Pediatrics, Henderson, Nevada; gDepartment of Pediatrics, University Hospital, Universitetssjukhuset MAS, Malmö, Sweden; hDepartment of Infectious Diseases, Malarsjukhuset, Eskilstuna, Sweden

Financial Disclosure: Drs Bernstein, Black, Conti, Twiggs, Flodmark, and Rombo and the Refrigerator-Stable Formulation Study Group for ProQuad participated in this research study, which was sponsored by Merck & Co, Inc. Drs Reisinger and Keyserling also participated in this research study and are speakers for Merck & Co, Inc.

ABSTRACT

OBJECTIVE. A refrigerator-stable formulation of ProQuad has been developed to expand the utility of ProQuad to areas in which maintenance of a frozen cold chain (−15°C or colder) during storage and transport may not be feasible. The objective of this study was to demonstrate that the immunogenicity and safety profiles of a refrigerator-stable formulation of ProQuad are similar to the recently licensed frozen formulation.

METHODS. In this double-blind, randomized, multicenter study, healthy 12- to 23-month-old children with negative vaccination and clinical histories for measles, mumps, rubella, varicella, and zoster were vaccinated with either the refrigerator-stable formulation of ProQuad (N = 1006) or the frozen formulation of ProQuad (N = 513). Patients were followed for 42 days after vaccination for adverse experiences. Immunogenicity was evaluated 6 weeks after vaccination.

RESULTS. The refrigerator-stable formulation of ProQuad was generally well tolerated. The incidence of adverse experiences was similar between groups. No vaccine-related serious adverse experiences were reported. For both groups, the response rate was ≥97.7% for measles, mumps, and rubella, and the percentage of patients with a varicella zoster virus antibody titer of ≥5 U/mL glycoprotein antigen-based enzyme-linked immunosorbent assay after vaccination was ≥88.8%. The geometric mean titers for all antigens were numerically slightly higher in patients who received the refrigerator-stable formulation.

CONCLUSIONS. The refrigerator-stable formulation of ProQuad is generally well tolerated, highly immunogenic, and noninferior in terms of postvaccination antibody responses. This refrigerator-stable formulation may improve ease of vaccine administration, increase use of the vaccine throughout the world because of its improved storage conditions, and replace the frozen formulation of ProQuad or any dose of M-M-RII and Varivax in routine practice.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2283
doi:10.1542/peds.2006-2283

Key Words: measles, mumps, rubella, varicella, vaccine, ProQuad, Varivax, M-M-RII, immunization

Abbreviations
VRC—vaccination report card
VZV—varicella zoster virus
ELISA—enzyme-linked immunosorbent assay
gpELISA—glycoprotein antigen–based enzyme-linked immunosorbent assay
GMT—geometric mean titer
CI—confidence interval

Accepted for publication Nov 21, 2006
Address correspondence to Henry H. Bernstein, DO, General Academic Pediatrics, Dartmouth Hitchcock Medical Center, I Medical Center Dr, Lebanon, NH 03756. E-mail: henry.bernstein@hitchcock.org

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
In the United States and other countries, M-M-RII (Merck & Co, Inc, West Point, PA) has been highly effective in reducing the incidence of measles, mumps, and rubella. To date, vaccination rates against measles, mumps, and rubella have reached 93% in the United States and are generally high in most developed countries. Routine use of Varivax (Merck & Co, Inc) in the United States has resulted in a substantial reduction in the incidence of varicella. However, vaccination rates against varicella have reached only 87.5% in the United States despite Advisory Committee on Immunization Practices recommendations in 1996 for universal use in young children. Varivax is not currently recommended in most other countries.

ProQuad is a combined measles, mumps, rubella, and varicella virus vaccine developed by Merck & Co, Inc. Concerns regarding an already complex vaccination schedule and additional costs to both parents and health care providers are notable obstacles surrounding vaccine administration. As a combination vaccine, ProQuad is expected to decrease the number of injections that are given to children and ultimately improve compliance and immunizations rates for the 4 diseases worldwide.

The frozen formulation of ProQuad, licensed in the United States and Europe, has been shown to be as immunogenic and generally well tolerated as its component vaccines, M-M-RII and Varivax, and is starting to be widely used in the United States. However, its utility is limited to geographic areas where a frozen cold chain (−15°C or colder) during transport and storage in clinics, pharmacies, and physicians’ offices can be maintained. Maintaining a refrigerated cold chain (2–8°C), as is done with most other vaccines, would be feasible in a much larger part of the world. Because of this, Merck & Co, Inc has developed a refrigerable-stable formulation of ProQuad, incorporating ≤2.5 mg of a urea-based stabilizer in each dose. The effect of this stabilizer on immunogenicity and safety profiles was previously evaluated in a clinical study of Varivax (varicella virus vaccine live [Oka/Merck]). Both Varivax formulations, with and without the stabilizer, were generally well tolerated, were comparable with respect to the occurrence of injection-site and systemic adverse experiences, and were highly immunogenic.

The purpose of this study was to demonstrate that the immunogenicity and the safety of a refrigerable-stable formulation of ProQuad are similar to the recently licensed frozen formulation. This study removes 1 barrier to effective vaccine delivery by lessening the burden for health care professionals of the storage and handling of frozen vaccines.

**METHODS**

**Study Patients**

This study was approved by the institutional review board/independent ethics committee of each of the 32 participating sites. The study sites consisted of academic institutions, regional and pediatric medical centers, managed care organizations, and private physician offices, with 30 sites in the United States and 2 sites in Sweden. The study was conducted from September 2002 through June 2003. Written informed consent was obtained from the parent/legal guardian of each patient before enrollment.

Healthy 12- to 23-month-old children with negative vaccination and clinical histories for measles, mumps, rubella, varicella, and zoster were eligible to participate. Patients were excluded for any of the following reasons: previously received measles, mumps, rubella, and/or varicella vaccine either alone or in combination; immunodeficient or receiving immunosuppressive therapy; history of a seizure disorder; known allergy to any vaccine component; recent exposure to measles, mumps, rubella, varicella, and/or zoster; receipt of any inactivated nonstudy vaccines within 14 days or live nonstudy vaccines within 30 days before enrollment; receipt of immune globulin or blood products within the 5 months before enrollment or had scheduled to receive such products within 42 days after vaccination; or, in the opinion of the investigator, had any condition that would have interfered with study objectives.

**Vaccine/Randomization**

Patients were randomly assigned at a 2:1 ratio to receive either the refrigerator-stable or the frozen formulation of ProQuad on day 1. All clinical materials were manufactured by Merck & Co, Inc. ProQuad is a sterile, lyophilized quadrivalent vaccine that after reconstitution is administered as a 0.5-mL subcutaneous injection. This study was double blind with regard to treatment group. The parent/legal guardian of the patient, the study personnel who administered the vaccine and handled all safety and serologic follow-up, and sponsor personnel were blinded to which vaccine formulation each patient received. Because of obvious differences in vaccine storage conditions and slight differences in appearance, the vaccines that were used in this study were prepared and accounted for by an unblinded third party who was otherwise not involved in the conduct of the study.

**Safety Surveillance**

Parents/legal guardians were asked to record on a vaccination report card (VRC) their child’s daily temperature (axillary), all local and systemic adverse experiences, and any other vaccines or medications administered both on the day of vaccination and for 41 additional days after vaccination. Parents/legal guardians were asked to notify study personnel immediately if their child experienced measles, a measles-like rash, rubella, a rubella-like rash, varicella, a varicella-like rash, zoster, a zoster-like rash, mumps, or mumps-like symptoms or if any serious adverse experience occurred. Se-
rious and vaccine-related adverse experiences were followed to resolution. Parents/legal guardians were instructed to measure the patient’s rectal temperature if the patient’s axillary temperature was ≥98.6°F (≥37.0°C). When a patient’s temperature was ≥102°F (≥38.9°C) oral or equivalent (≥101°F [≥38.3°C] axillary or ≥103°F [≥39.4°C] rectal), study personnel recorded the fever as an adverse experience. The parent/legal guardian was responsible for evaluating the maximum intensity of each adverse experience. Local adverse experiences of swelling and redness were evaluated by maximum size. The investigator assessed each reported adverse experience as to seriousness, action taken, and causal relationship to study vaccine.

**Laboratory Methods**
A 5- to 10-mL blood sample was collected immediately before and ~6 weeks after vaccination. To determine measles, mumps, rubella, and varicella zoster virus (VZV) antibody levels, serum samples were tested by Merck Research Laboratories (West Point, PA) using appropriately sensitive enzyme-linked immunosorbent assay (ELISA) methods for measles, mumps, and rubella and a glycoprotein antigen–based enzyme-linked immunosorbent assay (gpELISA) for VZV.**14–16** Serum samples with antibody levels <255 mIU/mL for measles, <10 antibody U for mumps, <10 IU/mL for rubella, and <1.25 gpELISA U/mL for VZV were considered to be seronegative. For measles, mumps, and rubella, the antibody response rates were defined as the proportion of patients who were seronegative before vaccination and became seropositive after vaccination. The antibody response rate for VZV was defined as the proportion of patients who were seronegative at baseline and whose postvaccination titer was ≥5 gpELISA U/mL. A VZV postvaccination titer ≥5 gpELISA U/mL has been previously shown to be highly correlated with long-term protection.17,18

**Statistical Methods**
The primary immunogenicity analysis was performed on a per-protocol basis. The analysis of antibody response rates to each vaccine antigen was based on the noninferiority test developed by Miettinen and Nurminen19 with study center stratification. Noninferiority of geometric mean titers (GMTs) was based on an analysis of variance model of log-adjusted titers for each antigen with study center stratification. To conclude noninferiority of the refrigerator-stable formulation of ProQuad (using a 1-sided .025 significance level), the response rates could be no more than 5 percentage points lower for measles, mumps, and rubella and no more than 10 percentage points lower for VZV. The GMTs could be no more than 1.5-fold lower for all antigens. In addition, it could be concluded that the refrigerator-stable formulation of ProQuad induced acceptable antibody response rates when the lower bound of the 95% confidence interval (CI) for each antigen was entirely above the prespecified criteria (90% for measles, mumps, and rubella and 76% for varicella). Success of the trial required satisfaction of all 3 immunogenicity hypotheses.

To address the primary hypothesis regarding safety, we compared the safety profiles of the refrigerator-stable and frozen formulations of ProQuad. For injection-site adverse experiences that occurred on days 1 to 5 after vaccination and for specific systemic clinical adverse experiences that occurred on days 1 to 42 after vaccination, risk differences were estimated, and the 95% 2-sided CI was provided. For adverse experiences that specifically were prompted for on the VRC, including injection-site redness, injection-site swelling, injection-site pain/tenderness, measles-like rashes, rubella-like rashes, varicella-like rashes, zoster-like rashes, and mumps-like symptoms, as well as the incidence of elevated temperature (defined as ≥102°F [≥38.9°C] oral equivalent), the corresponding P values were provided.

**RESULTS**

**Study Population**
A total of 1519 healthy children, 12 to 23 months of age, were vaccinated with either the refrigerator-stable formulation (N = 1006) or the frozen formulation (N = 513) of ProQuad. At vaccination, both groups were comparable with respect to gender, age, and race; ~50% of patients were female, and 73% were white. The mean age was 13.3 months at study entry.

**Safety**
No vaccine-related serious adverse experiences were reported within the follow-up period. The overall rate of serious adverse experiences between recipients of the refrigerated and frozen formulations of ProQuad was not statistically significant (0.7% and 0.4%, respectively). All serious adverse experiences were medical conditions that are generally expected of a pediatric population. The serious adverse experiences that were reported by the 7 patients who received the refrigerator-stable formulation were respiratory syncytial virus infection, pneumonia, dehydration, pharyngitis, gastroenteritis, and accidental exposure. The serious adverse experiences that were reported by the 2 patients who received the frozen formulation were gastroenteritis, pneumonia aspiration, and neuroblastoma. The patient with neuroblastoma did not receive the diagnosis until 5 days after vaccination and continued in the study with no additional adverse experiences reported.

Table 1 summarizes the specific injection-site adverse experiences that were reported days 1 to 5 after vaccination. Overall, the proportion of patients who reported injection-site adverse experiences was comparable between the treatment groups. Injection-site adverse ex-
Table 1: Injection-Site Adverse Experiences 1 to 5 Days After Administration of ProQuad

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Refrigerator-Stable (N = 1006), n (%)</th>
<th>Frozen (N = 513), n (%)</th>
<th>Risk Difference (Refrigerator-Stable − Frozen), % (95% CI)*</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients with ≥1 injection-site adverse experience</td>
<td>370 (37.6)</td>
<td>190 (38.0)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Injection-site erythema</td>
<td>175 (17.8)</td>
<td>90 (18.0)</td>
<td>−0.2 (−4.5 to 3.8)</td>
<td>0.93</td>
</tr>
<tr>
<td>Injection-site hemorrhage</td>
<td>15 (1.5)</td>
<td>6 (1.2)</td>
<td>0.3 (−1.2 to 1.5)</td>
<td>NA</td>
</tr>
<tr>
<td>Injection-site pain</td>
<td>291 (29.6)</td>
<td>152 (30.4)</td>
<td>−0.8 (−5.8 to 4.1)</td>
<td>0.75</td>
</tr>
<tr>
<td>Injection-site rash</td>
<td>3 (0.3)</td>
<td>3 (0.6)</td>
<td>−0.3 (−1.5 to 0.4)</td>
<td>0.40</td>
</tr>
<tr>
<td>Injection-site swelling</td>
<td>86 (8.7)</td>
<td>46 (9.2)</td>
<td>−0.5 (−3.7 to 2.5)</td>
<td>0.77</td>
</tr>
</tbody>
</table>

Percentages were calculated on the basis of the number of patients with follow-up. No risk differences, CIs, and/or P values were planned. N indicates number of patients vaccinated in each treatment group; n, number of patients with an adverse experience; NA, not applicable.

* Risk differences and CIs were based on the pooled incidence rates across study centers; corresponding values are provided only for events that were prompted for on the VRC and have been rounded to the nearest hundredth.

Table 2: Systemic Adverse Experiences (≥5% in Either Treatment Group) 1 to 42 Days After Administration of ProQuad

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Refrigerator-Stable (N = 1006), n (%)</th>
<th>Frozen (N = 513), n (%)</th>
<th>Risk Difference (Refrigerator-Stable − Frozen), % (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients with ≥1 systemic adverse experience</td>
<td>779 (79.2)</td>
<td>400 (80.0)</td>
<td>−0.9 (−5.3 to 3.4)</td>
</tr>
<tr>
<td>Upper respiratory tract infection</td>
<td>196 (19.3)</td>
<td>104 (20.8)</td>
<td>−1.4 (−5.8 to 2.8)</td>
</tr>
<tr>
<td>Otitis media/ear infectionb</td>
<td>185 (18.8)</td>
<td>101 (20.2)</td>
<td>−0.1 (−4.3 to 3.8)</td>
</tr>
<tr>
<td>Elevated temperaturec</td>
<td>168 (17.1)</td>
<td>86 (17.2)</td>
<td>−1.7 (−5.7 to 2.1)</td>
</tr>
<tr>
<td>Nasopharyngitis</td>
<td>139 (14.1)</td>
<td>79 (15.8)</td>
<td>−0.0 (−3.2 to 2.9)</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>110 (11.2)</td>
<td>46 (9.2)</td>
<td>2.0 (−1.4 to 5.1)</td>
</tr>
<tr>
<td>Dermatitis diaper</td>
<td>85 (8.6)</td>
<td>43 (8.6)</td>
<td>0.0 (−3.2 to 2.9)</td>
</tr>
<tr>
<td>Cough</td>
<td>78 (7.9)</td>
<td>40 (8.0)</td>
<td>−0.1 (−3.2 to 2.7)</td>
</tr>
<tr>
<td>Irritability</td>
<td>73 (7.4)</td>
<td>49 (9.8)</td>
<td>−2.4 (−5.7 to 0.6)</td>
</tr>
<tr>
<td>Rash (nonspecific)</td>
<td>72 (7.3)</td>
<td>38 (7.6)</td>
<td>−0.3 (−3.3 to 2.4)</td>
</tr>
<tr>
<td>Vomiting</td>
<td>67 (6.8)</td>
<td>29 (5.8)</td>
<td>1.0 (−1.8 to 3.5)</td>
</tr>
<tr>
<td>Rhinorhrea</td>
<td>65 (6.6)</td>
<td>33 (6.6)</td>
<td>0.0 (−2.9 to 2.6)</td>
</tr>
<tr>
<td>Rash morbilliform</td>
<td>48 (4.9)</td>
<td>32 (6.4)</td>
<td>−1.5 (−4.3 to 0.9)</td>
</tr>
</tbody>
</table>

Percentages were calculated on the basis of the number of patients with follow-up. There were no significant differences between the treatment groups for any adverse experiences shown above. N indicates number of patients vaccinated in each treatment group; n, number of patients with an adverse experience.

* Risk differences and CIs were based on the pooled incidence rates across all study centers.

b Otitis media and ear infection were reported as separate adverse experiences but are combined here because of similarity in terms. Otitis media was reported in 147 (15.0%) patients who received the refrigerator-stable formulation and in 76 (15.2%) patients who received the frozen formulation. Ear infection was reported in 38 (3.9%) patients who received the refrigerator-stable formulation and in 25 (5.0%) patients who received the frozen formulation.

c Includes any temperature ≥102°F (≥38.9°C), oral equivalent.
temic adverse experiences were compared, there were no significant differences between treatment groups.

The proportion of patients with elevated temperatures (≥102°F [≥38.9°C], oral equivalent) was comparable between treatment groups. More than 45% of fevers were reported during days 6 to 13 after vaccination. The majority of the fever episodes were classified as either mild or moderate and were of short duration.

The proportions of patients with a VRC-prompted rash (measles-like, rubella-like, and varicella-like) were comparable in both treatment groups, with no statistically significant differences observed. Mumps-like symptoms were also prompted for on the VRC, although none were reported. For both treatment groups, the percentage of patients who reported a measles-like, rubella-like, and varicella-like rash was ≤6.4%, ≤1.2%, and ≤3.0%, respectively.

**Immunogenicity**
Response rates and GMTs 6 weeks after vaccination for measles, mumps, rubella, and varicella are shown in Table 3. For both the refrigerator-stable and frozen formulations of ProQuad, the response rates for measles, mumps, and rubella were ≥97.7% and the percentage of patients who had a baseline VZV antibody titer <1.25 gpELISA U/mL and achieved a postvaccination VZV antibody titer ≥5 gpELISA U/mL was ≥88.8%. Because the lower bound of the 95% CI on the difference in response rates between treatment groups was ≤5 percentage points for measles, mumps, and rubella and ≤10 percentage points for VZV, the response rates to each vaccine antigen in patients who received the refrigerator-stable formulation of ProQuad were considered similar to the frozen formulation. The proportion of patients who had a baseline VZV titer <1.25 gpELISA U/mL and achieved a postvaccination VZV titer ≥1.25 gpELISA U/mL was 98.9% for those who received the refrigerator-stable formulation of ProQuad and 97.9% for those who received the frozen formulation of ProQuad (data not shown). The response rates that were induced by the refrigerator-stable formulation of ProQuad were found to be acceptable because the lower bound of the 95% CIs was >90% for measles, mumps, and rubella, and the percentage of patients with a postvaccination VZV titer ≤5 gpELISA U/mL was >76%.

Postvaccination GMTs were numerically slightly higher in patients who received the refrigerator-stable formulation of ProQuad; however, the difference was statistically <1.5-fold between the formulations for each vaccine antigen. Because the difference in GMTs was <1.5-fold, the refrigerator-stable formulation of ProQuad was considered similar to the frozen formulation. Success of the trial required that the statistical analysis show that the refrigerator-stable formulation induced both response rates and GMTs that were noninferior and that the response rates were entirely above the prespecified criteria (90% for measles, mumps, and rubella and 76% [≥5 gpELISA U/mL] for varicella).

**DISCUSSION**
The refrigerator-stable formulation of ProQuad was highly immunogenic in 12- to 23-month-old children and was generally well tolerated. No vaccine-related serious adverse experiences were reported, and there were no clinically significant differences in injection-site and systemic adverse experiences between the 2 formulations. These results were generally comparable to those reported in earlier trials with the frozen formulation, except for the rates of elevated temperature (≥102°F [≥38.9°C], oral equivalent), which were lower in this study. The lower rates were most likely attributed to a change in data collection methods in this study, including confirmation of elevated axillary temperatures by the rectal method and exclusion of qualitative reports of “warm to touch” as temperatures ≥102°F (≥38.9°C).

One limitation of this study was that the number of patients was not adequate to detect rare adverse experiences that occurred at a rate of 1 per 10,000 or less.

| TABLE 3 | Measles, Mumps, Rubella, and VZV Response Rates and GMTs 6 Weeks After Vaccination
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Antigen (Assay)</td>
<td>Parameter</td>
<td>Refrigerator-Stable (N = 1006)</td>
<td>ProQuad (N = 513)</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>Observed Response</td>
<td>95% CI</td>
</tr>
<tr>
<td>Measles (ELISA)</td>
<td>% ≥255 mIU/mL</td>
<td>879</td>
<td>99.1% (871/879)</td>
</tr>
<tr>
<td></td>
<td>GMT</td>
<td>2416</td>
<td>96.5%–98.6%</td>
</tr>
<tr>
<td>Mumps (ELISA)</td>
<td>% ≥10 antibody U</td>
<td>883</td>
<td>97.7% (863/883)</td>
</tr>
<tr>
<td></td>
<td>GMT</td>
<td>97</td>
<td>96%–99%</td>
</tr>
<tr>
<td>Rubella (ELISA)</td>
<td>% ≥10 IU/mL</td>
<td>908</td>
<td>99.6% (904/908)</td>
</tr>
<tr>
<td></td>
<td>GMT</td>
<td>97</td>
<td>96%–99%</td>
</tr>
<tr>
<td>Vancella (gpELISA)</td>
<td>% ≥5 gpELISA U/mL</td>
<td>839</td>
<td>90.1% (756/839)</td>
</tr>
</tbody>
</table>

N indicates number of patients vaccinated in each treatment group; n, number of patients with measles antibody titers <255 mIU/mL, mumps antibody titers <10 ELISA antibody U, rubella antibody titers <10 IU/mL, or VZV antibody titers <1.25 gpELISA U/mL at baseline and with postvaccination serology contributing to the per-protocol analysis.

a GMTs for measles, mumps, and rubella have been rounded to the nearest whole number.
Such rare adverse experiences may be studied best in routine postlicensure surveillance. Long-term, postlicensure, passive surveillance already provides extensive safety experience for the components of ProQuad (M-M-RII and Varivax).

In recent years, the recommended immunization schedule for children in the United States has become increasingly complex, with children receiving 20 or more injections in their first 18 months of life. The use of combination vaccines, such as ProQuad, is an effective way to overcome health care provider and parental concerns with multiple injections. In addition, combination vaccines may improve the timely administration of vaccines, increase vaccination coverage rates, reduce costs, and allow for the addition of new vaccines into the routine immunization schedule.20,21

Frozen vaccines add complexity to storage and handling in various settings. ProQuad is currently available only as a frozen formulation, and the need to store it in the freezer limits its utility outside the United States.21

World Health Organization guidelines for countries that are considering the implementation of a new vaccine that is intended for large-scale public health use mandate that the ability to overcome technical limitations such as vaccine storage temperature and physical storage capacity be considered.20,22 For many developing and remote areas of the world, successful implementation of a vaccine necessitates use of a product that does not require frozen storage.23,24 Accordingly, vaccines, with the exception of oral polio vaccine, that are supplied as part of the World Health Organization-sponsored Expanded Program on Immunization are recommended to be thermostable at refrigerated temperatures through the cold chain, originating with the manufacturer and ending with either regional/district vaccine stores or local health/daily use centers.25,26 Availability of the refrigerator-stable formulation of ProQuad, especially where the vaccine distribution infrastructure is not sufficient to manage frozen vaccine storage and handling, will further aid the reduction of measles, mumps, and rubella diseases and significantly decrease the incidence of varicella disease in countries other than the United States by facilitating the ease of routine immunization against varicella.

CONCLUSIONS
Adopting this refrigerator-stable formulation of ProQuad will lessen the burden of distribution and storage on pediatric practices, increase the ease of vaccine administration, and allow additional global expansion of current recommendations throughout the world. This combination vaccine with its improved storage conditions may also replace any separate dose of M-M-RII and Varivax for immunization against measles, mumps, rubella, and varicella. The refrigerator-stable formulation of ProQuad would be expected to sustain already high global vaccination rates against measles, mumps, and rubella and increase vaccination rates against varicella.

ACKNOWLEDGMENTS
Members of the Refrigerator-Stable Formulation Study Group for ProQuad include Wilson P. Andrews, MD, Jerry C. Bernstein, MD, Joseph Bertino, PharmD, Steven B. Black, MD, Stanley L. Block, MD, Louis Brine, MD, Kevin Browngoehl, MD, Ralph M. Conti, MD, Matthew Cox, MD, Robert Dracker, MD, Carl-Erik Floidmark, MD, Michael Gerber, MD, Gregory C. Gray, MD, J. Randy Hedgepeth, MD, Harry Keyserling, MD, Paul Lei, MD, Michael Levin, MD, Carl Lindgren, MD, Stephen Luber, MD, Edgardo Malacaman, MD, Colin Marchant, MD, Michelle Ogle, MD, Keith S. Reisinger, MD, Lars Rombo, MD, Shelly D. Senders, MD, Julie Shepard, MD, Douglas Short, MD, Malcolm Sperling, MD, Robert Stacks, MD, James Troutman, MD, and Jerry D. Twiggis, MD.

We acknowledge the housestaff, faculty, nurses, and administrative staff at each participating institution for participation and support of this study. The editorial assistance and meticulous attention to detail of Alyssa Scott is appreciated. We thank the children and their families for participation in this research project.

REFERENCES
12. Mendez R, Henderson F, Reisinger K, et al. Immunogenicity and safety of process upgrade varicella vaccine (PUVV) with new stabilizer as compared to PUVV with current stabilizer when administered concomitantly with M-M-RII in healthy children. Presented at the 22nd annual meeting of the European Society for Paediatric Infectious Diseases; May 26–28, 2004; Tampere, Finland
Do Skinfold Measurements Provide Additional Information to Body Mass Index in the Assessment of Body Fatness Among Children and Adolescents?

Zuguo Mei, MD*, Laurence M. Grummer-Strawn, PhD*, Jack Wang, MS*, John C. Thornton, PhD*, David S. Freedman, PhD*, Richard N. Pierson, Jr, MD*, William H. Dietz, MD, PhD*, Mary Horlick, MD*

*Division of Nutrition and Physical Activity, Centers for Disease Control and Prevention, Atlanta, Georgia; †Body Composition Unit, Department of Medicine, Obesity Research Center, St Luke’s-Roosevelt Hospital, New York, New York; ‡National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland

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

ABSTRACT

OBJECTIVES. The purpose of this work was to validate the performance of age- and gender-specific BMI, triceps, and subscapular skinfold for the classification of excess of body fat in children and adolescents and to examine how much additional information these 2 skinfold measurements provide to BMI-for-age.

METHODS. The receiver operating characteristic curve was used to characterize the sensitivity and specificity of these 3 indices in classifying excess body fat. Percentage of body fat was determined by dual-energy radiograph absorptiometry. Both ≥85th and ≥95th percentile of percentage of body fat were used to define excess body fat. Data from the New York Pediatric Rosetta Body Composition Project were examined (n = 1196; aged 5–18 years).

RESULTS. For children aged 5 to 18 years, BMI-for-age, triceps skinfold-for-age, and subscapular skinfold-for-age each performed equally well alone in the receiver operating characteristic curves in the identification of excess body fat defined by either the 85th or 95th percentile of percentage of body fat by dual-energy radiograph absorptiometry. However, if BMI-for-age was already known and was >95th percentile, the additional measurement of skinfolds did not significantly increase the sensitivity or specificity in the identification of excess body fat.

CONCLUSIONS. In contrast to the recommendations of expert panels, skinfold measurements do not seem to provide additional information about excess body fat beyond BMI-for-age alone if the BMI-for-age is >95th percentile.
**Total Body Fatness** can be measured by a variety of methods, including laboratory techniques, such as underwater weighing, dual-energy radiograph absorptiometry (DXA), total-body water, total-body electrical conductivity, total-body potassium, and air displacement plethysmography. However, most of these methods are limited to research because of their complexity and cost. Studies of DXA measurements in animals and humans have demonstrated that DXA scans accurately capture regional and total body composition and may constitute a new reference method. However, the most frequently used tools in public health evaluations and clinical screening are anthropometric-based measurements, such as skinfold thickness or circumference, or various height- and weight-based indices, such as weight-for-height and BMI.

BMI has been recommended and used worldwide to screen for overweight and obesity among both adults and adolescents, but as a measure of weight relative to height, it is only a proxy used to estimate body fatness. Body fatness also has been estimated from measurements of skinfold thicknesses, which correlate reasonably well with various laboratory estimates of body fatness. However, concerns have been expressed about the accuracy of this approach, because skinfold measurements are poorly reproducible, and, typically, only a few regional body sites are measured.

Expert panels have recommended the measurement of triceps and subscapular skinfold thickness as a component of the in-depth medical assessment for children and adolescents with age- and gender-specific BMI ≥95th percentile or BMI ≥30 (whichever was smaller) or BMI ≥85th percentile but <95th percentile with complications of obesity. However, this recommendation was based on clinical judgment rather than on the ability of skinfolds to predict body fat. No pediatric studies have systematically examined the improvement in the prediction of body fat provided by skinfold measures when BMI-for-age is known. In addition, this recommendation preceded the publication of the Centers for Disease Control and Prevention (CDC) 2000 gender-specific BMI-for-age reference values.

The purpose of this analysis is to validate the performance of age- and gender-specific BMI, triceps, and subscapular skinfold percentiles in identifying children and adolescents with excess body fat and to determine how much information these skinfold measurements can contribute to that provided by BMI-for-age alone.

**METHODS**

**Sample Design**

The potential study population consisted of 1208 healthy children and adolescents who were participants in the Pediatric Rosetta Study at St Luke’s-Roosevelt Hospital Center in New York (1995–2000), a cross-sectional study of pediatric body composition. Healthy volunteers aged 5 to 18 years were recruited in the New York City area through local newspaper notices, announcements at schools and activity centers, and word of mouth. The study protocol was approved by the institutional review board of St Luke’s-Roosevelt Hospital Center, and consent was obtained from each volunteer’s parent or guardian. When appropriate, assent was also obtained from each volunteer. Participants’ normal health status was confirmed by a medical history from a parent or guardian and a physical examination, both taken at the time of the body composition evaluation. The pubertal stage of subjects (as defined by the criteria of Tanner) was determined by physical examination by the pediatric endocrinologist or nurse in children <11 or 12 years and by self-assessment in older subjects. Details of the study design and methods have been published elsewhere. After excluding 12 children with missing height, weight, or skinfold measurements, we were left with 1196 subjects for this analysis.

**Body Composition**

All of the anthropometric measurements were made by specially trained laboratory technicians who measured subjects’ body height to the nearest 0.1 cm with a wall-mounted stadiometer (Holtain, Crosswell, Wales) and their weight to the nearest 0.1 kg using a balance-beam scale (Weight Tronix, New York, NY) with subjects wearing a hospital gown and foam slippers. They also measured subjects’ triceps and subscapular skinfold thicknesses to the nearest 1 mm with a Lange caliper (Beta Technology, Inc, Cambridge, MD); both measurements were taken on subjects’ right side in accordance with standard procedures. All of the anthropometric measurements were taken by 2 technicians with identical training and similar measurement precision. All of the intraclass (between technicians) correlation coefficients were ≥0.89.

Whole body DXA scans were performed using Lunar models DPX with pediatric software 3.8G and DPX-L with pediatric software 1.5G (GE Lunar Corporation, General Electric, Madison, WI) in accordance with the manufacturer’s instructions. Each scan provided estimates of subjects’ fat mass and fat-free mass in kilograms and percentage of body fat (%BF). The coefficient of variation (CV) for repeated measures of percentage of fat by total body DXA scan in pediatric subjects (aged 5–17 years) in the St Luke’s-Roosevelt Body Composition Laboratory is 2%. Each morning, before subjects were evaluated, technicians scanned an anthropomorphic spine phantom made up of calcium hydroxyapatite embedded in a 17.5 × 15 × 17.5 cm Lucite block with both DXA instruments for quality control. The phantom was also scanned immediately before and after all of the maintenance visits.
by the DXA system manufacturer. The measured phantom bone mineral density was stable throughout the study period at 1.166 to 1.196 g/cm². Methanol and water bottles (8-L volume), simulating fat and fat-free soft tissues, respectively, were scanned as soft-tissue quality control markers monthly; the range in measured R values over the study period was 1.255 to 1.258 (CV = 0.127%) for ethanol and 1.367 to 1.371 (CV = 0.103%) for water.

We used the age- and gender-specific BMI reference values developed by CDC29 to assign BMI percentiles and z scores (BMI-for-age) to study participants. To determine skinfold percentiles and z scores for study subjects, we developed triceps and subscapular skinfold references for children aged 2 to 18 years (triceps and subscapular skinfold-for-age) using the same CDC growth reference data set and smoothing techniques that were used to develop the BMI percentiles and z scores. To create an internal DXA %BF reference for excess body fat, we first divided the study participants by gender into 6 age groups based on the sample size (5–7, 8–9, 10–11, 12–13, 14–15, and 16–18 years), then calculated the

### Table 1: The 85th and 95th DXA Cutoff Points of %BF Reference for Excess Body Fat From the Pediatric Rosetta Body Composition Project Data Set

<table>
<thead>
<tr>
<th>Age, y</th>
<th>Boys n=626</th>
<th>Girls n=570</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>5–7</td>
<td>12.1 (3.5)</td>
<td>11.8 (3.4)</td>
<td>12.0 (3.4)</td>
</tr>
<tr>
<td>8–9</td>
<td>151.6 (19.8)</td>
<td>146.7 (16.2)</td>
<td>149.3 (18.3)</td>
</tr>
<tr>
<td>10–11</td>
<td>48.9 (19.6)</td>
<td>46.0 (17.9)</td>
<td>47.5 (18.9)</td>
</tr>
<tr>
<td>12–13</td>
<td>20.3 (4.5)</td>
<td>20.5 (4.9)</td>
<td>20.4 (4.7)</td>
</tr>
<tr>
<td>14–15</td>
<td>14.5 (8.9)</td>
<td>18.6 (9.3)</td>
<td>16.5 (9.3)</td>
</tr>
<tr>
<td>16–18</td>
<td>12.0 (9.0)</td>
<td>14.7 (9.9)</td>
<td>13.3 (9.5)</td>
</tr>
<tr>
<td>12.0 (10.6)</td>
<td>26.7 (9.9)</td>
<td>22.8 (10.0)</td>
<td></td>
</tr>
<tr>
<td>10.1 (8.7)</td>
<td>13.5 (9.6)</td>
<td>11.7 (9.3)</td>
<td></td>
</tr>
<tr>
<td>38.9 (14.9)</td>
<td>32.5 (9.8)</td>
<td>35.8 (13.1)</td>
<td></td>
</tr>
<tr>
<td>0.16 (0.99)</td>
<td>0.02 (1.08)</td>
<td>0.09 (1.04)</td>
<td></td>
</tr>
<tr>
<td>0.48 (1.09)</td>
<td>0.38 (1.14)</td>
<td>0.43 (1.12)</td>
<td></td>
</tr>
<tr>
<td>0.46 (1.10)</td>
<td>0.46 (1.07)</td>
<td>0.46 (1.09)</td>
<td></td>
</tr>
<tr>
<td>0.50 (1.06)</td>
<td>0.48 (1.12)</td>
<td>0.49 (1.09)</td>
<td></td>
</tr>
<tr>
<td>0.55 (0.89)</td>
<td>0.48 (0.93)</td>
<td>0.52 (0.91)</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2: Characteristics of and Body Composition Results for the Study Population (n = 1196)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Boys n=626</th>
<th>Girls n=570</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Height, cm</td>
<td>12.1 (3.5)</td>
<td>11.8 (3.4)</td>
<td>12.0 (3.4)</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>151.6 (19.8)</td>
<td>146.7 (16.2)</td>
<td>149.3 (18.3)</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td>48.9 (19.6)</td>
<td>46.0 (17.9)</td>
<td>47.5 (18.9)</td>
</tr>
<tr>
<td>Triceps skinfold, mm</td>
<td>20.3 (4.5)</td>
<td>20.5 (4.9)</td>
<td>20.4 (4.7)</td>
</tr>
<tr>
<td>Subscapular skinfold, mm</td>
<td>14.5 (8.9)</td>
<td>18.6 (9.3)</td>
<td>16.5 (9.3)</td>
</tr>
<tr>
<td>%BF</td>
<td>12.0 (9.0)</td>
<td>14.7 (9.9)</td>
<td>13.3 (9.5)</td>
</tr>
<tr>
<td>Fat mass, kg</td>
<td>19.2 (10.6)</td>
<td>26.7 (9.9)</td>
<td>22.8 (10.0)</td>
</tr>
<tr>
<td>Fat free mass, kg</td>
<td>10.1 (8.7)</td>
<td>13.5 (9.6)</td>
<td>11.7 (9.3)</td>
</tr>
<tr>
<td>Height-for-age z score</td>
<td>38.9 (14.9)</td>
<td>32.5 (9.8)</td>
<td>35.8 (13.1)</td>
</tr>
<tr>
<td>Weight-for-age z score</td>
<td>0.16 (0.99)</td>
<td>0.02 (1.08)</td>
<td>0.09 (1.04)</td>
</tr>
<tr>
<td>BMI-for-age z score</td>
<td>0.48 (1.09)</td>
<td>0.38 (1.14)</td>
<td>0.43 (1.12)</td>
</tr>
<tr>
<td>Triceps skinfold-for-age z score</td>
<td>0.46 (1.10)</td>
<td>0.46 (1.07)</td>
<td>0.46 (1.09)</td>
</tr>
<tr>
<td>Subscapular skinfold-for-age z score</td>
<td>0.50 (1.06)</td>
<td>0.48 (1.12)</td>
<td>0.49 (1.09)</td>
</tr>
</tbody>
</table>

### Table 3: Comparison of the Areas Under the ROC Curves (SEs in Parentheses) of BMI-for-Age, Triceps Skinfold-for-Age, and Subscapular Skinfold-for-Age to Identify Excess Fat Defined as %BF by DXA at the 85th or 95th Percentiles in Children 5 to 18 Years of Age

<table>
<thead>
<tr>
<th>Variable</th>
<th>85th Percentile Cutoff (SE)</th>
<th>95th Percentile Cutoff (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boys n=626</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI-for-age</td>
<td>0.952 (0.015)</td>
<td>0.966 (0.022)</td>
</tr>
<tr>
<td>Triceps skinfold-for-age</td>
<td>0.951 (0.015)</td>
<td>0.958 (0.024)</td>
</tr>
<tr>
<td>Subscapular skinfold-for-age</td>
<td>0.961 (0.014)</td>
<td>0.952 (0.026)</td>
</tr>
<tr>
<td>Girls n=570</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI-for-age</td>
<td>0.954 (0.016)</td>
<td>0.970 (0.021)</td>
</tr>
<tr>
<td>Triceps skinfold-for-age</td>
<td>0.952 (0.016)</td>
<td>0.955 (0.026)</td>
</tr>
<tr>
<td>Subscapular skinfold-for-age</td>
<td>0.947 (0.016)</td>
<td>0.962 (0.024)</td>
</tr>
<tr>
<td>Tanner stage 1 (n=425)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI-for-age</td>
<td>0.948 (0.020)</td>
<td>0.960 (0.029)</td>
</tr>
<tr>
<td>Triceps skinfold-for-age</td>
<td>0.947 (0.020)</td>
<td>0.947 (0.033)</td>
</tr>
<tr>
<td>Subscapular skinfold-for-age</td>
<td>0.958 (0.018)</td>
<td>0.939 (0.036)</td>
</tr>
<tr>
<td>Tanner stage 2–5 (n=740)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI-for-age</td>
<td>0.954 (0.014)</td>
<td>0.972 (0.018)</td>
</tr>
<tr>
<td>Triceps skinfold-for-age</td>
<td>0.953 (0.014)</td>
<td>0.965 (0.020)</td>
</tr>
<tr>
<td>Subscapular skinfold-for-age</td>
<td>0.950 (0.014)</td>
<td>0.969 (0.019)</td>
</tr>
</tbody>
</table>

### Explanatory Notes

85th and 95th percentiles for %BF from DXA by gender and age group. The 85th and 95th DXA cutoff points are presented in Table 1.

#### Statistical Methods

We used receiver operating characteristic (ROC) curves to assess the performance of the indices in identifying subjects with excess body fat as defined by the 85th and 95th percentile for %BF by DXA. We obtained the ROC curves by dichotomizing BMI-for-age and the skinfolds-for-age at all of the possible z score values and then plotting the resulting true-positive fractions (sensitivity)
on the y-axis versus the corresponding false-positive fractions (1 - specificity) on the x-axis. An area under the curve of 0.5 indicates that the test is no better than chance in identifying cases, whereas an area of 1.0 indicates perfect prediction. We used MedCalc software to test the significance of the differences for the areas under the ROC curves (AUC). First, using the 85th or 95th percentile for %BF from DXA as a standard, we examined the ROC performance of BMI-for-age, triceps skinfold-for-age, and subscapular skinfold-for-age measurements independently to evaluate which was the best indicator of excess body fat with the population stratified...
RESULTS

Subject characteristics are presented in Table 2. The study population was 52% boys, and the racial/ethnic distribution was 25% white, 22% black, 14% Hispanic, 30% Asian, and 8% other. The percentages of subjects in pubertal stages 1, 2, 3, 4, and 5 were 36%, 19%, 16%, 16%, and 13%, respectively, for boys and 38%, 11%, 17%, 19%, and 15%, respectively, for girls. Boys were taller and heavier ($P < .01$) and had higher $z$ scores for height than girls ($P < .05$), but the mean BMI values of boys and girls were similar. Girls had greater triceps and subscapular skinfold thicknesses and higher percentages of body fat and total fat mass, whereas boys had more fat-free mass ($P < .01$). However, the $z$ scores for skinfold thicknesses did not differ significantly between boys and girls.

Table 3 shows the ROC performance of BMI-for-age, triceps skinfold-for-age, and subscapular skinfold-for-age measurements among subjects stratified by age group, pubertal stage, gender, and race. There were no significant differences in the ROC performance of the 3 indicators in any of the stratified groups ($P > .05$). The areas under the curves were all $>0.91$.

Figure 1A compares the results obtained using ROC curves for BMI-for-age alone, triceps skinfold-for-age conditional on BMI $\geq$85th percentile, and triceps skinfold-for-age conditional on BMI $\geq$95th percentile to classify excess body fat in children and adolescents ages 5 to 18 years, using 95% for %BF by DXA as the standard. The key portion of the ROC curves from sensitivities $\geq70\%$ and specificities $\geq70\%$ ($1 - $specificities $\leq30\%$) is amplified in Fig 1B. Selected sensitivities and specificities of BMI-for-age alone from Fig 1 are listed in Table 4, and selected sensitivities and specificities for triceps skinfold-for-age conditional on BMI $\geq$85th and 95th percentiles are listed in Table 5.

For the BMI $\geq$95th percentile group, the triceps skinfold-for-age (solid line with star markers in Fig 1B) does not provide useful information, because, whereas it improves the specificity for excess body fat compared with BMI-for-age alone at the 95%, the increase is actually less than would be found by simply increasing the BMI percentile cutoff. However, for the BMI $\geq$85th percentile group, triceps skinfold-for-age $\geq$95th percentile adds information to that from BMI-for-age alone by increasing the specificity for excess body fat (solid line with diamond markers in Fig 1B and Table 5). Results from the same analysis repeated using subscapular skinfold-for-age (shown in Fig 2 and Table 5) were quite similar.

### Table 4

**Selected Sensitivities and Specificities of BMI-for-Age**

<table>
<thead>
<tr>
<th>BMI Percentile</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>85th</td>
<td>98.5</td>
<td>72.7</td>
</tr>
<tr>
<td>90th</td>
<td>97.0</td>
<td>79.7</td>
</tr>
<tr>
<td>95th</td>
<td>92.5</td>
<td>89.4</td>
</tr>
<tr>
<td>96th</td>
<td>91.0</td>
<td>91.7</td>
</tr>
<tr>
<td>97th</td>
<td>89.6</td>
<td>94.8</td>
</tr>
<tr>
<td>98th</td>
<td>74.6</td>
<td>96.9</td>
</tr>
<tr>
<td>99th</td>
<td>46.3</td>
<td>99.4</td>
</tr>
</tbody>
</table>

### Table 5

**Selected Sensitivities and Specificities of Triceps Skinfold-for-Age and Subscapular Skinfold-for-Age**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Triceps Skinfold</th>
<th>Subscapular Skinfold</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sensitivity, %</td>
<td>Specificity, %</td>
</tr>
<tr>
<td>BMI $\geq$ 95th percentile + skinfold $\geq$ 75th percentile</td>
<td>92.5</td>
<td>89.5</td>
</tr>
<tr>
<td>BMI $\geq$ 95th percentile + skinfold $\geq$ 85th percentile</td>
<td>92.5</td>
<td>89.9</td>
</tr>
<tr>
<td>BMI $\geq$ 95th percentile + skinfold $\geq$ 90th percentile</td>
<td>91.0</td>
<td>90.8</td>
</tr>
<tr>
<td>BMI $\geq$ 95th percentile + skinfold $\geq$ 95th percentile</td>
<td>89.6</td>
<td>93.2</td>
</tr>
<tr>
<td>BMI $\geq$ 95th percentile + skinfold $\geq$ 99th percentile</td>
<td>47.8</td>
<td>98.4</td>
</tr>
<tr>
<td>BMI $\geq$ 85th percentile + skinfold $\geq$ 75th percentile</td>
<td>98.5</td>
<td>75.1</td>
</tr>
<tr>
<td>BMI $\geq$ 85th percentile + skinfold $\geq$ 85th percentile</td>
<td>98.5</td>
<td>78.0</td>
</tr>
<tr>
<td>BMI $\geq$ 85th percentile + skinfold $\geq$ 90th percentile</td>
<td>97.1</td>
<td>82.4</td>
</tr>
<tr>
<td>BMI $\geq$ 85th percentile + skinfold $\geq$ 95th percentile</td>
<td>92.5</td>
<td>89.3</td>
</tr>
<tr>
<td>BMI $\geq$ 85th percentile + skinfold $\geq$ 99th percentile</td>
<td>47.8</td>
<td>98.3</td>
</tr>
</tbody>
</table>
as were results from the same analyses in which the 85th percentile for %BF by DXA was used to define excess body fat (results not shown).

**DISCUSSION**

Because of the limitations of BMI as a measure of actual body fatness for overweight or obesity, expert panels have recommended measuring triceps and subscapular skinfold thicknesses as part of the in-depth medical assessment of children and adolescents with age- and gender-specific BMI ≥95th percentile or ≥30 (which ever was smaller) or age- and gender-specific BMI ≥85th percentile but <95th percentile or equal to 30 (which ever was smaller).\(^\text{15,16}\) However, the additional informa-
tion provided by these skinfolds has not been rigorously examined. Our results indicate that for children aged 5 to 18 years, BMI-for-age, triceps skinfold-for-age, and subscapular skinfold-for-age individually performed equally well in the classification of excess body fat defined by total body DXA. However, if BMI-for-age is already known and is >95th percentile, the specificity in the identification of excess fat is similar for BMI-for-age alone and in combination with skinfold-for-age. On the other hand, for subjects with BMI-for-age of 85th to 95th percentile and skinfold-for-age ≥95th percentile, the specificity for identification of excess body fat is slightly improved.

Our results are based on standardized measurements of height, weight, and triceps and subscapular skinfold thickness on >1000 children and adolescents aged 5 to 18 years, as well as on DXA measurements for these children. Although the children and adolescents recruited in this study were from a convenience sample and, thus, do not represent the general US population, their mean height, weight, and BMI were only slightly different from children and adolescents examined in the National Health and Nutrition Examination Surveys at a fairly comparable time period. In our analysis, we assigned age- and gender-specific percentiles/z scores for BMI and for triceps and subscapular skinfold thicknesses and then examined the sensitivity and specificity of their ROC performance in the classification of excess body fat using %BF by DXA as the reference.

Skinfold measurements have been widely used to assess body composition in the past. They are simpler and less expensive than hydrostatic weighing or other laboratory-based techniques for body composition analysis. After the outlay for purchase of calipers, the costs are minimal. However, measurement can vary from tester to tester depending on skill and experience, unless cross-validation between testers and test-retest reliability evaluation are performed and monitored. On the other hand, BMI is calculated from weight and height measurements, both of which are routinely performed in pediatric clinical settings and are more reliable than skinfold thickness measurements. The procedure of measuring height and weight is simpler, the reliability is higher compared with that of measuring skinfolds, and both are routinely performed in pediatric clinical settings. Because BMI is calculated from weight and height, and reference curves for BMI-for-age are readily available, it is an appropriate screening test for excess adiposity. Furthermore, more recent studies concluded that BMI is an excellent proxy measure of adiposity in children and adolescents. Our results show that >90% of children were correctly classified as having high or low body fat using BMI-for-age ≥95th percentile cutoff (Table 4).

CONCLUSIONS

This study provides evidence that the current expert panel recommendation to measure the triceps and subscapular skinfolds as part of the in-depth assessment of pediatric patients with BMI-for-age ≥95th percentile to confirm excess body fat may need to be reevaluated. This recommendation preceded the publication of the CDC 2000 gender-specific BMI-for-age reference. Using these reference data and triceps and subscapular skinfold references based on the same data set, our evaluation suggests that skinfold measurements provide additional information on excess body fat for pediatric subjects ages 5 to 18 years with BMI-for-age of the 85th to 95th percentiles but not for those with BMI-for-age >95th percentile.

ACKNOWLEDGMENT

This work was supported by National Institutes of Health grant DK37352.

REFERENCES

Perinatal Circulating Visfatin Levels in Intrauterine Growth Restriction

Ariadne Malamitsi-Puchner, MD; Despina D. Briana, MD; Maria Boutsikou, MD; Evangelia Kouskouni, MD; Demetrios Hassiakos, MD; Demetrios Gourgiotis, PhD

*Neonatal Division, Second Department of Obstetrics and Gynecology, and Research Laboratories, Second Department of Pediatrics, Athens University, Athens, Greece

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

ABSTRACT

OBJECTIVE. The objective of this study was to investigate possible alterations in circulating levels of the adipocytokine visfatin in intrauterine growth-restricted and normal pregnancies, given that these groups differ considerably in fetal nutrition, body fat mass, and metabolic/endocrine mechanisms.

METHODS. Serum visfatin levels were prospectively measured by enzyme immunoassay in 40 mothers and their 40 singleton term fetuses and neonates on postnatal days 1 and 4. Twenty neonates had intrauterine growth restriction (birth weight ≤3rd customized centile, adjusted for parameters that influence growth potential), and 20 were appropriate for gestational age.

RESULTS. Circulating maternal visfatin levels were significantly elevated in pregnancies with intrauterine growth restriction compared with control pregnancies with appropriate-for-gestational-age infants and negatively correlated with customized centiles in the group with intrauterine growth restriction. Postnatal day-1 and -4 visfatin levels were significantly higher in neonates with intrauterine growth restriction compared with neonates who were appropriate for gestational age. Postnatal-day-1 prefeeding insulin levels were significantly lower in neonates with intrauterine growth restriction.

CONCLUSIONS. Pathologic conditions in pregnancy that lead to intrauterine growth restriction could be responsible for elevated maternal visfatin levels. Higher visfatin levels in neonates with intrauterine growth restriction may serve as an early marker with prognostic value for later development of insulin resistance or type 2 diabetes, whereas lower insulin levels may indicate reduced β-cell mass and/or impaired β-cell function.
Insulin resistance, obesity-related diabetes, and accompanying metabolic disorders are strongly associated with increased visceral adipose tissue mass.\(^1\)\(^-\)\(^4\) Visfatin, a 52-kd visceral fat-specific adipokine,\(^5\) probably links the expansion of adipose depot to insulin resistance.\(^6\) Visfatin, which is identical to pre-B-cell colony-enhancing factor (PBCEF),\(^7\) is immunolocalized in both normal and infected human fetal membranes\(^8\) and upregulated during labor.\(^9\)

Newborns with asymmetric intrauterine growth restriction (IUGR)\(^10\)\(^-\)\(^12\) are at increased risk for development of metabolic syndrome later in life,\(^11\)\(^-\)\(^14\) as a result of insulin resistance.\(^11\)\(^-\)\(^17\) We hypothesized that circulating visfatin levels should differ between infants with IUGR and control infants who are appropriate for gestational age (AGA), because the former present reduced fat mass\(^18\)\(^-\)\(^19\) and undergo adaptational changes of endocrine/metabolic mechanisms as a result of intrauterine malnutrition.\(^11\) Therefore, we aimed to evaluate circulating visfatin concentrations in IUGR and AGA mother-infant pairs at crucial perinatal time points.

**METHODS**

The ethics committee of our teaching hospital approved the study protocol. Participating mothers provided signed informed consent before enrollment. Forty parturients who gave birth consecutively either to 20 term singleton infants who were AGA or to 20 term singleton infants who had asymmetric IUGR (birth weight ≤3rd customized centile) were recruited. The Gestation Related Optimal Weight computer-generated program\(^20\)\(^,\)\(^21\) was used to calculate the customized centile for each pregnancy. Significant determinants of birth weight (maternal height and booking weight, ethnicity, parity, gestational age, and gender) were entered to adjust the normal birth weight centile limits.\(^20\)

Possible causes of IUGR were in 9 cases preeclampsia and in 11 cases pregnancy-induced hypertension plus various pathologic conditions (iron-deficient anemia \([n = 3]\), gestational diabetes \([n = 2]\), hypothyroidism \([n = 3]\), extreme obesity \([n = 2]\), and cardiac arrhythmias \([n = 1]\)) and smoking >10 cigarettes per day \((n = 5)\).

Doppler studies (pulsatility index \([PI]\)) were performed in the IUGR group every 10 to 15 days, starting from the 32nd gestational week. Concerning uterine and umbilical arteries, \(PI\) values were in the upper physiologic limits for gestational age in 13 cases, whereas in 7, they showed increased impedance to flow. \(PI\) values of middle cerebral arteries were in the lower physiologic limits for gestational age, indicating initiation of blood flow redistribution process. Amniotic fluid was diminished in all IUGR cases. For its evaluation, the largest fluid column on the vertical plane was assessed and defined as diminished when <2 cm. Placental weight ranged from 255 to 400 g.

In the AGA group, mothers were healthy nonsmokers. Ultrasound studies were evaluated as nonpathologic, and placentas were normal in appearance and weight.

Tests for congenital infections were negative in all women of both groups, and neonates showed no symptoms of intrauterine infection or signs of genetic syndromes. One- and 5-minute Apgar scores were ≥8 in all neonates.

The demographic data of participating mothers and infants are listed in Table 1. Blood was collected in pyrogen-free tubes from mothers during the first stage of labor or before anesthesia in cases of elective cesarean section, from doubly-clamped umbilical cords, reflecting fetal state, and from neonates on postnatal days 1 (N1) and 4 (N4). Blood was immediately centrifuged after clotting, and supernatant serum was kept frozen at −80°C until assay. Determination of visfatin levels was performed by enzyme immunoassay (visfatin C-terminal [human] EIA; Phoenix Pharmaceuticals, Belmont, CA). Minimum detectable concentration and intraassay and interassay coefficients of variation were 0.1 ng/mL and 5% and 12%, respectively.

Visfatin data, in contrast to insulin data, were normally distributed (Kolmogorov-Smirnov test). The effect of various parameters (maternal age, customized centile, birth weight, mode of delivery, gender, and parity) on circulating visfatin levels was assessed using linear regression analysis. Nonparametric procedures (Mann-Whitney U test and Wilcoxon rank-sum test), \(\chi^2\) test, and Pearson’s or Spearman’s rank correlation coefficient, where appropriate, were applied. The SPSS 10.0 (SPSS, Chicago, IL) statistical software package was used.

**Table 1** Demographic Data of Participating Mothers and Their Fetuses/Neonates Who Were AGA or Had IUGR

| Table 1: Demographic Data of Participating Mothers and Their Fetuses/Neonates Who Were AGA or Had IUGR
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
</tr>
<tr>
<td>Birth weight, mean (SD), g</td>
</tr>
<tr>
<td>Customized centile, mean (SD)</td>
</tr>
<tr>
<td>Maternal age, mean (SD), y</td>
</tr>
<tr>
<td>Gestational age, mean (SD), wk</td>
</tr>
<tr>
<td>Male gender, (n)%</td>
</tr>
<tr>
<td>Female gender, (n)%</td>
</tr>
<tr>
<td>Vaginal delivery, (n)%</td>
</tr>
<tr>
<td>Cesarean section, (n)%</td>
</tr>
<tr>
<td>Parity, (n)%</td>
</tr>
</tbody>
</table>

NS indicates nonsignificant; NA, not applicable.
RESULTS
Determined mean (95% confidence intervals [CI]) values of circulating visfatin levels in both groups are shown in Fig 1. Maternal visfatin levels were significantly elevated (by 3.346 ng/mL on average) in the IUGR compared with the AGA group (regression coefficient $\beta$: 3.346 [95% CI: 0.049–6.644; $P = .047$]). Although infant customized centile was not an independent predicting variable for maternal visfatin levels, a significant negative correlation was found between the 2 variables in the IUGR group ($r = -0.411 [P = .008]$).

No significant differences in visfatin levels were demonstrated between fetuses with IUGR and fetuses that were AGA. N1 and N4 circulating visfatin levels were significantly higher in neonates who had IUGR than in neonates who were AGA (regression coefficient $\beta$: 2.80 [95% CI: 1.205–6.305; $P = .005$] and $\beta$: 2.769 [95% CI: 0.519–5.019; $P = .017$], respectively). Maternal and fetal visfatin levels were significantly correlated (AGA group $r = 0.742 [P < .001]$ and IUGR group $r = 0.478 [P = .033]$).

N1 prefeeding insulin levels (mean ± SD) were significantly lower in neonates who had IUGR than in neonates who were AGA (2.46 ± 3.57 vs 3.63 ± 2.80 $\mu$U/mL, respectively; $P = .048$). In contrast, maternal and N4 prefeeding insulin values did not vary significantly between IUGR and AGA groups (11.1 ± 12.18 vs 7.6 ± 5.21 and 4.04 ± 5.59 vs 4.58 ± 4.22 $\mu$U/mL, respectively). Circulating visfatin levels did not depend on gestational age, mode of delivery, or gender. Finally, no significant correlations were observed between serum visfatin and insulin levels in either group.

DISCUSSION
To our knowledge, this preliminary study is the first to determine circulating visfatin concentrations in neonates with IUGR. Although included pregnancies were limited, circulating visfatin levels were investigated in the same mother-infant pairs at 4 crucial perinatal time points, and statistically significant results were found between the IUGR and AGA groups.

Circulating maternal visfatin levels were higher in the IUGR than in the AGA group and negatively correlated with the customized centiles in the former. Although preeclampsia and pregnancy-induced hypertension, main causes of IUGR22 in our study, are associated with obesity23 and subsequently with increased visfatin levels,3 the latter could be related to other yet-unidentified mechanisms. Indeed, maternal pathologic conditions that are associated with IUGR22 (eg, protein and energy restriction,24-26 iron-deficient anemia27) have been reported in animal models to cause changes in maternal circulating levels of several hormones and adipokines, such as leptin.25 Therefore, altered maternal endocrine environment in pregnancies that are complicated with IUGR25 could be responsible for higher visfatin levels.

In addition, higher circulating visfatin levels were found in neonates who had IUGR. There are 2 potential explanations for this finding. The first relies on the fact that visceral adipose tissue is the predominant source of visfatin,3 and data suggest that newborns with low birth weight may have increased visceral fat stores.28,29 In contrast, Harrington et al,19 by applying MRI, found no differences in intra-abdominal adipose tissue between newborns who had IUGR and newborns who were AGA. In this study, fat mass was not directly measured and centrality of fat distribution was not assessed. Therefore, we could not document whether a relationship between circulating visfatin levels and visceral fat exists at birth.

A second explanation might involve the increasing evidence that circulating visfatin levels are higher in states of insulin resistance.5-30-34 Epidemiologic studies point to a strong relationship between IUGR and the development of insulin resistance and type 2 diabetes,35-37 the exact onset of which is not fully elucidated. Nevertheless, several studies,37,38-40 including this one, have determined lower insulin levels in neonates with IUGR (possibly as a result of reduced $\beta$-cell mass and poor intrauterine $\beta$-cell function40,41) and higher insulin sensitivity.17,38-40 Lévy-Marchal et al13 speculated that altered fetal development of adiposity in IUGR might permanently change the regulation of its metabolic and hormonal functions, predisposing to the later development of insulin resistance. Respectively, animal and human studies42-44 indicated that infants who are small for gestational age compared with control infants who are AGA gain more abdominal fat and body adiposity during...
postnatal life. Therefore, although higher visfatin levels in neonates with IUGR cannot be attributed to insulin resistance at birth, they could probably serve as an early marker with prognostic value for the later development of metabolic syndrome in this population. A similar hypothesis was previously made for circulating adiponectin levels in newborns who were small for gestational age.45

Maternal and fetal blood visfatin concentrations were strongly correlated in both groups, indicating the likelihood of transplacental transfer of visfatin. Finally, no significant differences in fetal visfatin levels were demonstrated between IUGR and AGA groups. Because visfatin/PBEF originates from human fetal membranes during pregnancy and is upregulated during labor, the lack of difference could be attributed to varying visfatin membrane production in the 2 groups. In addition, the lack of insulin/visfatin correlation may be attributable to the possibility that perinatal visfatin could mostly reflect PBEF. However, relevant information does not exist, and additional studies are needed to clarify these results.

CONCLUSIONS
Pathologic conditions in pregnancies that lead to IUGR may account for the elevated maternal visfatin levels. Higher visfatin levels in neonates with IUGR may serve as an early marker of insulin resistance or type 2 diabetes later in life. However, lacking adequate information on the physiologic role of visfatin in adults, it is difficult to speculate on its importance in the perinatal period. The source and the regulation of visfatin in the fetus and the neonate remain to be elucidated.

REFERENCES
subjects are reduced after gastric banding. J Clin Endocrinol Metab. 2006;91:1578–1581
34. Tan BK, Chen J, Digby JE, Keay SD, Kennedy CR, Randeva HS. Increased visfatin messenger ribonucleic acid and protein levels in adipose tissue and adipocytes in women with polycystic ovary syndrome: parallel increase in plasma visfatin. J Clin Endocrinol Metab. 2006;90:5022–5028
42. Ibanez L, Ong K, Dunger DB, de Zegher F. Early development of adiposity and insulin resistance after catch-up weight gain in small-for-gestational-age children. J Clin Endocrinol Metab. 2006;91:2153–2158
Differences in Severity-Adjusted Pediatric Hospitalization Rates Are Associated With Race/Ethnicity

James M. Chamberlain, MD<sup>a,b</sup>, Jill G. Joseph, MD, PhD<sup>c</sup>, Kantilal M. Patel, PhD<sup>a-d,†</sup>, Murray M. Pollack, MD, MBA<sup>a-g</sup>

<sup>a</sup>Department of Pediatrics, George Washington University School of Medicine, Washington, DC; Divisions of <sup>b</sup>Emergency Medicine and <sup>c</sup>Critical Care Medicine, Children’s National Medical Center, Washington, DC; <sup>d</sup>Department of Pediatrics, University of California School of Medicine, Davis, California; <sup>e</sup>Center for Health Services and Community Research, Washington, DC; <sup>f</sup>Children’s Research Institute, Washington, DC; <sup>g</sup>Center for Hospital-Based Specialties, Washington, DC

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

ABSTRACT

OBJECTIVE. Racial/ethnic disparities in health care delivery have been well described, but little is known about such disparities for children who seek emergency care. The objective of this study was to test the hypothesis that severity-adjusted emergency department pediatric admission rates are associated with race/ethnicity.

METHODS. Secondary analyses were conducted of an established database of 16 emergency departments that participated in a national study to validate the Pediatric Risk of Admission II score, which is used to measure severity of illness. Patients were randomly selected by the coordinating center from daily emergency department visit logs. Crude and severity-adjusted admission rates were compared among the 3 most common races/ethnicities: white, black, and Hispanic. Adjusted admission rates were calculated by using the standardized admission ratio, which was calculated by dividing the observed admissions by the predicted admissions, when predicted was calculated from the Pediatric Risk of Admission II score.

RESULTS. After exclusion of 3 sites that recorded race/ethnicity in <10% of patients, there were 13 sites with 8952 patients in the 3 major race/ethnicity groups. Black and Hispanic patients were similar to each other and different from white patients; therefore, these 2 groups were combined for analyses. Both crude (8.2% vs 5.3%) and severity-adjusted (standardized admission ratio: 1.71 vs 1.1) admission rates were higher in white than in nonwhite patients. Standardized admission ratios were close to 1.0 in both race/ethnicity groups in the higher quintiles of illness severity. In contrast, white patients were admitted at 1.5 to 2 times the expected rate in the lowest 2 quintiles of severity.

CONCLUSIONS. There are differences in both crude and adjusted admission rates between white and black/Hispanic patients. The results are more consistent with high rates of discretionary admissions for white patients with low illness severity than with underadmitting severely ill black or Hispanic patients.
inequities in health care on the basis of race and/or ethnicity are a serious quality problem facing the American health care system. In addition to obvious differences in access that are associated with social class and poverty, several authors have raised concerns that racial disparities in health care are not simply an issue of inadequate access. In several adult studies, even when care is available and adjustments are made for insurance status and income, medical care is unequal when racial/ethnic groups are compared. These disparities occur across a wide range of diseases and health circumstances, ranging from primary care preventive and screening services to care for end-stage renal disease, cancer, and HIV. Importantly, some studies have also demonstrated higher mortality rates in minorities associated with these disparities. In the emergency department (ED) setting, where access is guaranteed, there are significant racial differences in the care of patients with acute myocardial infarction and black and Hispanic patients are less likely to receive analgesic medications for bone fractures. In children, appendicitis is more likely to result in rupture and abscess in nonwhite compared with white patients.

With the exception of previous work done by our group, no studies have examined racial disparities in the treatment of children with emergency conditions. An important barrier to this type of research has been the lack of a method for adjusting for illness severity. We developed and previously described a severity of illness adjustment method called the Pediatric Risk of Admission (PRISA) score. PRISA uses the risk for hospital admission as an index of severity and was developed and tested in a single institution. PRISA was found accurate in an independent study in a single hospital in Canada and was recalibrated in 5 academic medical centers in the United States. The second-generation score (PRISA II) was redeveloped and validated in a national sample of 16 EDs. We conducted analyses that are reported here to test the hypothesis that there are differences in severity-adjusted rates of pediatric emergency hospitalization associated with race/ethnicity.

METHODS

The results presented in this article represent secondary analyses of data that were collected for the purpose of developing the PRISA II severity-of-illness score for pediatric emergency patients. The details of site and patient selection are described in detail elsewhere. In brief, 16 EDs were block randomly selected in a balanced 2 × 3 factorial design to represent the care factors of high or low volume (less than or equal to or more than the median of pediatric visits per year), presence or absence of a pediatric emergency medicine (PEM) subspecialist, and presence or absence of residents delivering ED care. Two sites were selected randomly for each stratum, resulting in 16 study sites. There were 8 sites with high volume and 8 with low volume, 8 sites with a PEM subspecialist and 8 without, and 8 sites with residents and 8 without residents.

The coordinating center randomly enrolled 2 patients per day from each ED using a random-number generator to select patients from each site’s daily consecutive patient arrival log. Data were collected for 375 consecutive days, but actual record collection ranged from 729 to 753 patients. For sites with >729, patients were randomly eliminated to keep only 729 patients from each site, thereby maintaining a balanced 2 × 3 factorial design.

Masked, photocopied data were abstracted and computerized. Abstracted data included demographic, historic, physiologic, diagnostic, and therapeutic variables. Race/ethnicity were recorded with the following categories, as commonly used by the participating hospitals: white, black, Hispanic, Asian, and other. There were too few patients with Asian or “other” for meaningful analysis; therefore, we compared outcomes among the 3 major categories of white, black, and Hispanic. Preliminary analyses revealed obvious and consistent differences between white patients and the other 2 groups but little difference between black and Hispanic patients. Therefore, apart from initial descriptive data, we compared white patients with the combination of black and Hispanic patients, reasoning that (1) both of these can be considered minorities when compared with white individuals, and (2) the observed differences were in the same direction and of approximately the same magnitude. Socioeconomic data were not available on individual patients; therefore, we used postal zip code–level median household income from the 2000 US census.

The primary outcome was hospital admission, which was defined as admission to an inpatient unit or to an observation area for >12 hours. Mandatory admissions were defined as admissions that received a therapy (eg, intravenous antibiotics) or had a condition (eg, full-thickness burn) that generally required inpatient care. The list of characteristics that defined mandatory admission was validated using a consensus process and has been published previously. All classifications of patients regarding admission and mandatory admission status were made by the principal investigator after review of inpatient records. Analyses used the previously developed PRISA II score, which models the risk for medically mandatory hospital admission using historical, physiologic, and therapeutic variables in a multiple logistic regression model. The PRISA II score was developed using generalized estimating equations (GEEs), an extension of generalized linear model regression analysis that incorporates the correlated data or effects of patient clustering within institutions.

For the analyses reported here, crude and adjusted admission rates were calculated and compared among the 2 race groups described. Adjusted admission rates were calculated using the standardized admission ratio.
(SAR; observed admissions divided by predicted mandatory admissions, where number of predicted mandatory admissions was calculated from the PRISA II model). A SAR of 1.0 indicates that the total number of admissions equals the predicted number of mandatory admissions on the basis of illness severity. A SAR >1 indicates admission rates in excess of those predicted by the severity of illness of the group, whereas SARs <1 indicate fewer admissions than predicted by severity. In addition, we used multivariable logistic regression modeling that included demographic variables to determine the additional effect of demographic severity on admission probability, expressed as adjusted odds ratios (ORs). Included in this model were the PRISA II score (illness severity), gender, race, and median household income. Household income was determined from the 2000 US census using the patient’s zip code. The GEE method was used to adjust for the clustering of patients within institutions. GEEs are an extension of logistic regression analysis that adjust the confidence intervals of the parameter estimates on the basis of the degree of clustering of patients. For example, if most of the Hispanic patients were treated at only a few institutions, then GEE would account for how much of the variance in admission rates was attributable to different admission practices in those institutions as well as how much of the variance is related to differences in the way Hispanic patients were treated compared with white patients. We did not include age as a separate variable because age is incorporated into the PRISA II score, both as a risk factor (age <90 days) and in the determination of reference ranges of physiologic variables.

RESULTS
A total of 11,664 pediatric patients were enrolled, 8668 (74.3%) of whom had 1 of the 3 major race categories recorded (white, black, and Hispanic). Three hospitals recorded race in <10% of patients; these sites were therefore excluded from additional analysis. Within the remaining 13 sites, 525 patients had other race categories (Asian, other, no race recorded) and were also excluded from additional analysis. Thus, the total sample for analysis included 8952 patients from 13 institutions (Fig 1).

The 3 major race/ethnicity groups were compared with respect to patient characteristics and the most common diagnoses (minor injuries, otitis media, and fever), as shown in Table 1. Black children were more likely to arrive by emergency medical services. Hispanic patients had higher baseline severity (ie, PRISA II scores) compared with the other groups. White children were more likely to present with minor injuries and less likely to have otitis media than black and Hispanic children.

Crude and adjusted admission rates by race are presented in Table 2. White patients had higher overall rates of admission and higher rates of admissions to ICUs. The difference in overall rates of admission persisted even after adjustment for illness severity using the SAR (Table 2 and Fig 2). Figure 3 depicts the SAR for the race/ethnicity groups for each quintile of predicted admission risk. It is apparent that as the severity of illness quintile rises, admissions are increasingly equivalent and close to predicted. However, in the 2 lowest risk quintiles, white patients are admitted at 1.5 to 2 times the rate predicted by the need for therapies or their physiologic status, which is not true for black and Hispanic patients, for whom the SAR is close to 1 in all quintiles. Table 3 depicts the GEE model (multiple logistic regression model), which incorporates demographic variables, including race, gender, median household income, and severity-of-illness score (PRISA II). The adjusted OR for admission for black and Hispanic children is approxi-
mately half (0.54) that of the reference group of white children.

**DISCUSSION**

These results are consistent with a growing body of literature demonstrating disparities in health care that are associated with race. Both crude and severity-adjusted admission rates were lower for black and Hispanic children when compared with white children. In multivariate analyses, the probability of hospital admission for black and Hispanic children was lower even after controlling for illness severity and sociodemographic characteristics. Therefore, in this patient sample, the effect of race/ethnicity on admission probability is independent of income measured at the zip code level.

Three important points about these results are worth noting. First, more health care is not always better. These data are more consistent with a practice of overadmitting white patients who are less severely ill than underadmitting black and Hispanic patients who are more severely ill. As depicted in Fig 3, white patients who are not severely ill are admitted in excess of predicted. It is in such less seriously ill patients that there is greater physician discretion in admission decision-mak-

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total Sample (N = 8952)</th>
<th>White (n = 3112 [34.8%])</th>
<th>Black (n = 3288 [36.7%])</th>
<th>Hispanic (n = 2552 [28.5%])</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD), y</td>
<td>6.8 (5.8)</td>
<td>7.4 (5.9)</td>
<td>7.1 (5.9)</td>
<td>5.4 (5.2)</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td>Male gender, %</td>
<td>4669 (52.2)</td>
<td>1672 (53.3)</td>
<td>1697 (51.6)</td>
<td>1300 (50.9)</td>
<td>.083*</td>
</tr>
<tr>
<td>Arrival by EMS, %</td>
<td>875 (9.8)</td>
<td>313 (9.5)</td>
<td>410 (13.2)</td>
<td>265 (10.4)</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td>PRISAI score, mean (SD)</td>
<td>1.68 (4.09)</td>
<td>1.51 (4.21)</td>
<td>1.35 (4.02)</td>
<td>2.05 (4.01)</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td>Minor injury, %c</td>
<td>2089 (23.3)</td>
<td>950 (30.5)</td>
<td>680 (20.7)</td>
<td>459 (18.0)</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td>Otitis media, %d</td>
<td>866 (9.7)</td>
<td>237 (7.2)</td>
<td>330 (10.6)</td>
<td>299 (11.7)</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td>Fever, %</td>
<td>439 (4.9)</td>
<td>180 (5.7)</td>
<td>173 (5.3)</td>
<td>86 (3.4)</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td>Median household income (zip code level), mean (SD), $</td>
<td>37 027 (15 789)</td>
<td>46 152 (19 147)</td>
<td>33 029 (9949)</td>
<td>31 030 (11 894)</td>
<td>&lt;.001*</td>
</tr>
</tbody>
</table>

EMS indicates emergency medical services.  
* Kruskal-Wallis test.  
* x² test.  
* Strain, sprain, abrasion, laceration, or fracture without associated cranial, thoracic, or abdominal injury.  
* In patients with normal immune status.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Total Sample (N = 8952)</th>
<th>White (n = 3112 [34.8%])</th>
<th>Black (n = 3288 [36.7%])</th>
<th>Hispanic (n = 2552 [28.5%])</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Admission rate, %</td>
<td>564 (6.30)</td>
<td>254 (8.16)</td>
<td>176 (5.35)</td>
<td>134 (5.25)</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td>ICU admissions, %</td>
<td>71 (0.79)</td>
<td>41 (1.32)</td>
<td>22 (0.67)</td>
<td>8 (0.31)</td>
<td>&lt;.001*</td>
</tr>
<tr>
<td>SAR ≤ SE</td>
<td>564/431.3 (1.308 ± 0.042)</td>
<td>254/148.7 (1.708 ± 0.071)</td>
<td>176/152.2 (1.156 ± 0.070)</td>
<td>134/130.4 (1.028 ± 0.077)</td>
<td>&lt;.001bc</td>
</tr>
</tbody>
</table>

The SAR was calculated by dividing the observed number of admissions by the predicted number of mandatory admissions, as predicted by the PRISAI II score. Patients with mandatory admissions received an inpatient therapy or had a physiologic status that generally required inpatient care (see “Methods”).  
* x² test.  
* Hosmer-Lemeshow goodness-of-fit test.  
* P < .001, white versus black and white versus Hispanic; P = .217, black versus Hispanic.

**FIGURE 2**  
SAR according to race/ethnicity. The SAR was calculated by dividing the observed number of admissions by the predicted number of mandatory admissions, as predicted by the PRISAI II score. * P < .001 comparing white patients to black patients and white patients to Hispanic patients.
In contrast, for both white and nonwhite patients, admission rates are very close to predicted in the more severely ill quintiles. These results are consistent with a national study of myringotomy in preschool children. White patients were more likely to receive a myringotomy,27 a surgical procedure for which there are clear medical indications in only 41% of patients.28 The second point is the critical importance of accounting for illness severity when studying health care disparities. Without the ability to measure severity quintiles, as displayed in Fig 3, we might have erroneously concluded that black and Hispanic patients were being denied essential hospital admissions. Instead, our results suggest that white patients are overadmitted when not severely ill. Third, the magnitude of the effects of race/ethnicity that were observed in this study (adjusted OR: 0.54; Table 3) is consistent with previous literature, which demonstrated adjusted ORs of 0.47 to 0.60.3,8,9,12

The appropriate disposition of patients to either inpatient care or outpatient care is one of the most important decisions made by ED physicians. Admitting patients unnecessarily creates iatrogenic risk29,30 and unnecessarily increases health care costs,29 whereas failing to admit patients who require hospitalization may allow progression of disease without adequate therapy. Therefore, we chose to study the association of hospital admission with race/ethnicity in this large national database of ED visits. One significant advantage of this study is the ability to control for illness severity, including physiologic variables. The PRISA II score was developed and validated using this patient sample.25 Therefore, we were able to assess the independent contribution of race/ethnicity to admission probability after controlling for physiologic and other indicators of illness severity.

There are several limitations to this study. First, race/ethnicity was recorded on medical charts by various personnel, including physicians, nurses, and clerical staff. Race is a social construct rather than a biological trait; therefore, the accuracy of assignment is not really at issue. The results of this study suggest that however race is defined in EDs, it is associated with differential probability of hospitalization. Second, we did not collect insurance information or household income on individual patients, and there are well-described limitations of using postal code median household income as a proxy for individual income.31 Finally, there are many immeasurable factors in the complex decision to hospitalize a child, including parental preferences and community and family resources. Given our study design, it is impossible to determine the reasons for the observed differences in treatment. Language barriers, for example, might be associated with admission decisions for Hispanic patients but would be unlikely to affect black patients. Despite these limitations, the results reported here may be the first to identify a racial/ethnic differential in pediatric ED disposition independent of physiologic and medical factors. Additional studies are required to explore possible explanations for the relationship be-

FIGURE 3
SARs according to predicted admission risk quintile. The SAR was calculated by dividing the observed number of admissions by the predicted number of mandatory admissions, as predicted by the PRISA II score.

TABLE 3  GEE Model Incorporating Severity Score (PRISA II), Race, Income, and Gender

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>Adjusted OR (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>−3.2020</td>
<td>0.3560</td>
<td>1.00 (reference)</td>
<td></td>
</tr>
<tr>
<td>PRISA II score</td>
<td>0.2342</td>
<td>0.0091</td>
<td>1.26 (1.24–1.29)</td>
<td>.0001</td>
</tr>
<tr>
<td>Race/ethnicity (white)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race/ethnicity (black or Hispanic)</td>
<td>0.6130</td>
<td>0.2453</td>
<td>0.54 (0.33–0.88)</td>
<td>.013</td>
</tr>
<tr>
<td>Median household income</td>
<td>0.0000</td>
<td>0.0000</td>
<td>1.00</td>
<td>818</td>
</tr>
<tr>
<td>Gender</td>
<td>0.1016</td>
<td>0.0878</td>
<td>1.11 (0.93–1.31)</td>
<td>.247</td>
</tr>
</tbody>
</table>

CI indicates confidence interval.
between ED physician decision-making and patient race/ethnicity.

ACKNOWLEDGMENTS

This study was supported by Agency for Healthcare Quality and Research grant RO1 HS10238-02. The agency acted solely as the funding body for this investigator-initiated research and did not have input into the study design, conduct of the study, data analysis, or generation of the manuscript.

This article is dedicated to the memory of Dr Patel.

REFERENCES

The Spectrum of Cardiac Anomalies in Noonan Syndrome as a Result of Mutations in the *PTPN11* Gene

Yves Sznajer, MDa, Boris Keren, MDa, Clarisse Baumann, MDa, Sabrina Pereiraa, Corinne Alberti, PhDp, Jacques Elion, MD, PhDp, Hélène Cavé, MDa, Alain Verloes, MD, PhDp,c

aDepartment of Medical Genetics, bClinical Epidemiology Unit, and cInstitut National de la Santé et de la Recherche Médicale INSERM U676, AP-HP Robert Debré University Hospital, Paris, France

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

**ABSTRACT**

**OBJECTIVE.** Noonan syndrome is a clinically homogeneous but genetically heterogeneous condition. Type 1 Noonan syndrome is defined by the presence of a mutation in the *PTPN11* gene, which is found in ~40% of the cases. Phenotype descriptions and cardiac defects from cohorts with Noonan syndrome were delineated in the “pregenomic era.” We report the heart defects and links to gene dysfunction in cardiac development in a large cohort of patients with type 1 Noonan syndrome.

**METHODS.** This was a retrospective, multicenter study based on clinical history, pictures, and medical and cardiologic workup over time. Data were collected by referral geneticists. Mutation screening was performed by direct sequencing of exons 2, 3, 4, 7, 8, 12, and 13 and their intron-exon boundaries, which harbor 98% of identified mutations the *PTPN11* gene.

**RESULTS.** A *PTPN11* gene mutation was identified in 104 (38.25%) of 274 patients with Noonan syndrome. Heart defect was present in 85%. The most prevalent congenital heart defects were pulmonary valve stenosis (60%), atrial septal defect, ostium secundum type (25%), and stenosis of the peripheral pulmonary arteries (in at least 15%). Pulmonary valve stenosis and atrial septal defect, ostium secundum type, were significantly associated with the identification of a mutation in the *PTPN11* gene. Ventricular septal defect and most left-sided heart defects showed a trend toward overrepresentation in the group without a mutation.

**CONCLUSION.** We compared our data with previous series and integrated the comprehension of molecular *PTPN11* gene dysfunction in heart development.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-0211
doi:10.1542/peds.2006-0211

Dr. Sznajer’s current affiliation is Clinical Genetics Unit, Hôpital Universitaire des Enfants Reine-Fabiola, Université Libre de Bruxelles, Brussels, Belgium.

**Key Words**
Noonan syndrome, cardiac defect, *PTPN11* gene, phenotype, genotyping

**Abbreviations**
NS—Noonan syndrome
CHD—congenital heart defect
PVS—pulmonary valve stenosis
HCM—hypertrophic cardiomyopathy
ASD-OS—atrial septal defect of ostium secundum type
VSD—ventricular septal defect
RAS—rat sarcoma viral oncogene homolog
MAPK—mitogen-activated protein kinase
Egfr—epithelial growth factor receptor
NFAT—nuclear factor of activated T cell
LEOPARD—lentigines, EKG anomalies, ocular hypertelorism, pulmonary stenosis, abnormal genitalia, deafness

Accepted for publication Dec 12, 2006

Address correspondence to Alain Verloes, MD, PhD, Clinical Genetics Unit, 48 Boulevard Séruzier, 75019 Paris Cedex 19, France. E-mail: alain.verloes@rdb.ap-hop-paris.fr

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275. Copyright © 2007 by the American Academy of Pediatrics
NOONAN SYNDROME (NS) is an autosomal dominant multiple congenital anomalies syndrome that is characterized by minor facial anomalies (hypertelorism, downward slant of the palpebral fissures with ptosis, and low-set posteriorly angulated ears), short stature, congenital heart defects (CHDs), and variable learning disabilities. With an estimated incidence of 1 in 2000,1 NS turned out to be 1 of the most common genetic conditions in pediatrics and the second most common syndromal form of CHD, after trisomy 21.2 Type 1 NS corresponds to the causally homogeneous subgroup of patients who have NS and are molecularly defined by the presence of a mutation in the PTPN11 gene. Mutations in PTPN11 are responsible for ~40% of the cases of NS and for >95% of patients with LEOPARD syndrome ( lentigines, EKG anomalies, ocular hypertelorism, pulmonary stenosis, abnormal genitalia, deafness). The penetrance of the mutations of PTPN11 is 100%, but their expressivity is variable even within families. Very recently, mutations in KRAS3 and SOS14,5 were shown in, respectively, <5% and 15% of PTPN11-negative patients with NS. PTPN11, SOS1, and KRAS are part of the RAS signaling pathway. The causative gene(s) remains to be found for roughly half of patients with NS.

A broad spectrum of cardiac phenotypes has been recognized in NS. Pulmonary valve stenosis (PVS), hypertrophic cardiomyopathy (HCM) and atrial septal defect of ostium secundum type (ASD-OS) are the most common defects in NS, but ventricular septal defect (VSD), peripheral pulmonary stenosis, atrioventricular canal, aortic stenosis, mitral regurgitation, aortic coarctation, and coronary anomalies can also be seen in NS. Previous studies of cohorts of patients with NS suggested that the prevalence of specific types of CHD was different in type 1 NS and in NS without PTPN11 mutation.6–11 We report the cardiac anomalies in a large cohort of patients with type 1 NS to delineate the spectrum of cardiac defects that are caused by PTPN11 mutations and better understand the role of PTPN11 in heart development.

METHODS

Patients with NS or LEOPARD syndrome were recruited through a multicenter network of clinical geneticists from France, Switzerland, and Belgium for molecular analysis of the PTPN11 gene. The patients were entered systematically in the study, on the basis of the order by which samples arrived in the laboratory for diagnosis. Checklists were systematically sent to referring geneticists independently and before completion of the DNA screening. Call for supplementary data was not based on the identification (or not) of a mutation. Specific cardiological data were derived from records of transthoracic echocardiography and, when available, from cardiac catheterization and/or heart surgery. Because the data were mainly retrospective, distinction between PVS with normal valves and pulmonary valve dysplasia was hampered by the impossibility to check retrospectively for the presence of valvular dysplasia in patients for whom only pulmonary stenosis was mentioned. For this reason, we decided to merge the 2 anomalies under the term PVS.

Mutation screening focused on hot-spot locations in the PTPN11 gene, which covered 98% of identified mutations.6–9 Exons 2, 3, 4, 7, 8, 12, and 13 of the PTPN11 gene and their intron-exon boundaries were sequenced on both strands from genomic DNA for each patient on an ABI 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA). Mutations were identified using ABI PRISM SeqScape 2.0 (Applied Biosystems) by comparison with the reference sequence for genomic and complementary DNA sequence (GenBank accession Nos. NT.009775.14 and NM.002834).

Statistical analysis for qualitative data and comparison used χ² and exact Fisher test. Significance was determined as P < .05. Informed consent was obtained for inclusion and DNA screening. The ethics committee of AP-HP Bichat University Hospital (Paris, France) approved the study.

RESULTS

A total of 291 patients with NS or LEOPARD syndrome were initially enrolled in the study. The 19 patients who presented with café-au-lait spots and/or lentigines were found to share a small number of specific mutations. Because their clinical and molecular diagnosis was LEOPARD syndrome, they were studied and published separately.12 The final cohort of patients with NS comprises 149 boys and 123 girls. The mean age in the cohort is 7 years for the groups both with and without mutation (lower quartile: 2 years; upper quartile: 13 years).

A mutation in the PTPN11 gene was identified in 104 (38.25%) patients. Table 1 summarizes the cardiologic findings in patients. The c.922A→G (p.N308D) nucleotide substitution represents the most prevalent mutation (18.25%) in our patients. Table 2 presents the spectrum of CHD by functional domain, by exon, and, specifically, by those observed with the p.N308D mutation.

Among patients with a mutation, the apparent rate of heart defect is 85%. Because the selection criteria for offering molecular testing were clinically based and potentially biased toward overrepresentation of CHD-bearing patients, the true penetrance of CHDs in type 1 NS can nevertheless not be established objectively. The most prevalent CHD in type 1 NS are PVS (60%), ASD-OS (25%), and stenosis of the peripheral pulmonary arteries (in at least 15.4%). PVS and ASD-OS are significantly associated with the presence of a mutation (P < .01), and coarctation of aorta and HCM are associated with the absence of mutation. VSD and most left-sided heart defects show a trend toward overrepresentation in the group without a mutation. Comparison with previously reported series of genotyped NS is difficult. We summa-
rized the published data on CHDs in type 1 NS in Table 3 and compared the relative incidence of major CHDs in type 1 NS versus the general population based, for example, on prevalence data from Eurocat Database registries from European Western countries and/or Italian and/or Iceland databases (Table 4).

DISCUSSION

PTPN11 is a gene that is located on chromosome band 12q24.1 and encodes the nonreceptor protein tyrosine phosphatase SHP-2. The protein contains a phosphatase domain (PTP domain) and 2 amino-terminal SH2 domains (N-SH2 and C-SH2) that selectively bind to short amino acid motifs that contain a phosphotyrosyl residue. These SH2 domains promote SHP-2 association with cell surface receptors and cell adhesion molecules and modulate the cells’ responses to extracellular signals by regulating the phosphorysine content of specific intracellular proteins. More specific, SHP-2 functions as an intracellular enhancer of signal transduction for several growth factors, hormones, and cytokine receptors that are involved in the RAS/MAPK (rat sarcoma viral oncogene homolog/mitogen-activated protein kinase) signaling pathway, although its precise role in the pathway, as inhibitor of RAS inactivation by GAP (guanosine-triphosphatase–activating protein) proteins or as enhancer of the kinase cascade beyond RAS activation, remains elusive.13

The N-SH2 domain is a conformational switch: it either binds to and inhibits the PTP domain or binds phosphotyrosine proteins and allows the PTP catalytic domain to bind to specific phosphoproteins residues and hydrolyze them.8 Almost all reported mutations so far are missense exon changes, except for a unique in-frame 3-bp deletion.14 Almost all NS mutations are gain-of-function mutations that seem to alter N-SH2–PTP interaction, turning protein into a constitutionally active state.15–17 Some of them might further affect SH2 domain-phosphopeptide affinity and/or substrate specificity.18 Mutations cluster in 7 exons (2, 3, 4, 7, 8, 12, and 13). Those exons encode the interacting portions of SH2 and the PTP domains or the hinges segment binding the 3 domains. Exons 3 and 8 harbor 80% of the mutations. A single mutation (c.922A>G-p.N308D-exon 8) is present in ~20% of patients. Contrasting with NS mutations, LEOPARD syndrome mutants are catalytically defective and act as dominant negative mutations.16,19,20

The spectrum of CHDs has been widely described in

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Comparison of CHD Frequency Between PTPN11-Positive (Type 1 NS) and PTPN11-Negative Patients in Our Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHD</td>
<td>Patient With an Identified PTPN11 Mutation (n = 104, n (%)</td>
</tr>
<tr>
<td>PVS</td>
<td>62 (59.60)</td>
</tr>
<tr>
<td>ASD</td>
<td>26 (26.90)</td>
</tr>
<tr>
<td>Atroventricular canal</td>
<td>2 (1.90)</td>
</tr>
<tr>
<td>HCM</td>
<td>12 (10.60)</td>
</tr>
<tr>
<td>Coarctation of aorta</td>
<td>1 (0.96)</td>
</tr>
<tr>
<td>VSD</td>
<td>7 (6.70)</td>
</tr>
<tr>
<td>Aortic stenosis</td>
<td>2 (1.90)</td>
</tr>
<tr>
<td>Mitral insufficiency</td>
<td>5 (4.81)</td>
</tr>
<tr>
<td>Peripheral pulmonary artery stenosis</td>
<td>16 (15.40)</td>
</tr>
</tbody>
</table>

NS indicates not significant.

<table>
<thead>
<tr>
<th>TABLE 2</th>
<th>Spectrum of CHDs Reported With the 3 Functional Domains of the SHP2 Protein and Within the 7 Screened Exons</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHD</td>
<td>N-SH2 Domain</td>
</tr>
<tr>
<td>----------------</td>
<td>---------------</td>
</tr>
<tr>
<td>All CHD</td>
<td>All 2 3</td>
</tr>
<tr>
<td>PVS</td>
<td>27 (43.55)</td>
</tr>
<tr>
<td>ASD</td>
<td>13 (50.00)</td>
</tr>
<tr>
<td>HCM</td>
<td>5 (18.50)</td>
</tr>
<tr>
<td>Coarctation of aorta</td>
<td>0 0</td>
</tr>
<tr>
<td>VSD</td>
<td>0 0</td>
</tr>
<tr>
<td>Aortic stenosis</td>
<td>0 0</td>
</tr>
<tr>
<td>Mitral insufficiency</td>
<td>5 (18.50)</td>
</tr>
</tbody>
</table>

The last column summarizes the CHD associated with the single common N308D mutation located on exon 8.
NS before the era of *PTPN11* gene identification: PVS, ASD-OS, VSD, Ebstein anomaly, and coarctation of aorta were the most common heart defects encountered in patients with NS, as well as HCM. Among 151 patients who had NS and were aged 1 week to 45 years, Sharland et al reported PVS in 63%, HCM in 17%, and septal defects in 6%. Heart was normal in 10%. In a series of 184 patients, Digilio et al underlined that coarctation of aorta was present in 8.7%. In a cohort of 157 children, Marino et al emphasized high prevalence of atrioventricular canal (15.4%), whereas PVS represented only 38.9%. HCM 9.5%, coarctation of aorta 8.8%, ASD-OS 8%, mitral valve anomaly 5.8%, tetralogy of Fallot 4.4%, VSD 4.4%, patent ductus arteriosus 2.2%, and pulmonary atresia 1.4%.

Since the discovery of *PTPN11* involvement in type 1 NS, 6 small cohorts of patients have been reported from Europe and Asia. The spectrum of CHD in type 1 NS cannot be drawn easily from these cohorts (Table 3). In the genotype/phenotype cohort study by Tartaglia et al, only PVS, ASD-OS, and HCM were considered. PVS was present in 70.6% of the group with mutation and in 46.2% of the group without mutation (*P* < .01), whereas HCM was more prevalent in patients without a mutation (26% vs 5.9%; *P* < .01). The other CHDs were not reported and a “*PTPN11*-negative” group was not described. In a study of 96 patients with NS, Musante et al found a mutation rate of 29%. PVS was present in 60%; no HCM was identified among *PTPN11*-positive patients. There was no information on cardiac findings in the group without mutation, and only PVS, septal defects (of all types), and HCM were reported. The cardiac defects in the *PTPN11*-positive group were limited to the presence (or absence) of PVS, ASD-OS, VSD, and HCM. Sarkozy et al reported 73 patients with NS or LS and identified a mutation in 37% of the cases. Among patients with a mutation, PVS, HCM, partial AVC, and ASD-OS were the most common anomalies. No coarctation of aorta, tetralogy of Fallot, or VSD was observed.

A trend for ASD-OS to be correlated with exon 3 mutations was suggested. All mutations in the N-SH2 domain were associated with the presence of a heart defect, whereas in 70%, the mutations occurred in the PTP domain. Zenker et al reported 57 children with NS. The mutation rate was 60%. PVS was noted in 88% and HCM was noted in 3% of *PTPN11*-positive patients. The cardiac findings in the group without mutation were not given. Among 45 cases, Yoshida et al detected PVS in 55% of patients with mutation. As in Tartaglia’s and Musante’s studies, only PVS, HCM, or ASD-OS was recorded. In the most recent study, only PVS, HCM, ASD-OS were described separately.

Preselection criteria may explain differences between our study and the previous ones. Tartaglia et al included patients with clinical phenotype “evocative” of NS and PVS or HCM and/or pectus deformity or electrocardio-

### Table 3

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of cases</td>
<td>54</td>
<td>23</td>
<td>32</td>
<td>34</td>
<td>18</td>
<td>76</td>
<td>104</td>
<td>104 (100)</td>
</tr>
<tr>
<td>No. of patients with cardiac defect</td>
<td>42</td>
<td>15</td>
<td>20</td>
<td>20</td>
<td>9</td>
<td>40</td>
<td>34</td>
<td>34 (100)</td>
</tr>
<tr>
<td>PVS</td>
<td>36</td>
<td>13</td>
<td>21</td>
<td>30</td>
<td>10</td>
<td>38</td>
<td>62</td>
<td>62 (100)</td>
</tr>
<tr>
<td>ASD-OS</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>9</td>
<td>17</td>
<td>17 (100)</td>
</tr>
<tr>
<td>VSD</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>9</td>
<td>17</td>
<td>17 (100)</td>
</tr>
<tr>
<td>HCM</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>7</td>
<td>7 (100)</td>
</tr>
<tr>
<td>Coarctation of aorta</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Atrioventricular canal</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NM (0)</td>
</tr>
<tr>
<td>Mitral insufficiency</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NM (0)</td>
</tr>
<tr>
<td>Pulmonary branch stenosis</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NM (0)</td>
</tr>
</tbody>
</table>

**NM** indicates not mentioned; **NP** not possible (missing data).
gram anomaly, whereas Zenker et al9 used a combination of 2 of 3 criteria: typical heart defect (PVS or HCM), typical craniofacial morphology, and pterygium colli. The last suggested that his higher mutation rate was attributable to more stringent and rigorous inclusion criteria. These criteria may bias to selecting patients with the 2 most prevalent defects. No data have been published on the incidence of peripheral pulmonary stenosis in NS type 1. Physiologically, transient peripheral pulmonary stenosis is encountered in normal newborns (during transition from fetal to postnatal). Its incidence in our series is unlikely to reflect a physiologic transient state.

Comparison of patients with and without mutation is presented in Table 1. PVS was identified in 59.6% among mutation carriers (62 of 104) and 31.5% in the noncarrier group (53 of 168), ASD-OS respectively in 25% (26 of 104) and 13% (22 of 168), and stenosis of the peripheral pulmonary arteries in 15.4% (16 of 104) and 10.1% (17 of 168). Left-sided involvement was associated with absence of an identified PTPN11 mutation; mitral insufficiency in 4.8% (5 of 104) vs 7.1% (12 of 168), aortic stenosis in 1.9% (2 of 104) vs 4.7% (8 of 168), and coarctation of aorta in 0.96% (1 of 104) vs 5.95% (10 of 168; \( P < .01 \)). Finally, presence of VSD occurred in 6.75% (7 of 104) vs 13.7% (23 of 168; \( P < .001 \)).

Genotype-phenotype correlations within type 1 NS have been briefly addressed. In our cohort, nucleotide substitutions occurred in 44 (42.30%) in the N-SHP domain (exons 2 and 3), 5 (4.8%) in the C-SHP domain, 3 (2.9%) at the N-SHP/C-SHP hinge region, and 52 (50%) in the PTP domain. Exon 2 mutations were identified in <1%. Exon 3 mutations occurred in half of the cases; PVS was associated in 46% (29 of 62) of the cases. Exons 4 and 7 mutations were observed in 4.8% and 5.8%, respectively. Exon 8 mutations occurred in 26%. The N308D substitution has an incidence of 18.26% (19 of 104). PVS is the more common cardiac defect associated with exon 8 mutations, accounting for 58% (11 of

---

**TABLE 4**  Cardiac Anomalies Identified in Our Patients With Type 1 NS (in %)

<table>
<thead>
<tr>
<th>Cardiac Anomaly</th>
<th>Prevalence (per 10,000), Atlanta, GA, 1995–199719</th>
<th>Italian Multicentric Study 1992–1993</th>
<th>Iceland Study 1990–1999</th>
<th>Type 1 NS, %</th>
<th>Relative Risk</th>
<th>Eurocat Prevalence Database 1999–2003 (per 10,000 births)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCM</td>
<td>NM</td>
<td>NM</td>
<td>10.58</td>
<td>NA</td>
<td>NM</td>
<td></td>
</tr>
<tr>
<td>Heterotaxias and L-TGA</td>
<td>1.60</td>
<td>3.70</td>
<td>0.00</td>
<td>0.40</td>
<td>3.13</td>
<td></td>
</tr>
<tr>
<td>Outflow tract defects, total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tetralogy of Fallot</td>
<td>4.70</td>
<td>3.30</td>
<td>3.00</td>
<td>0.00</td>
<td>0.25</td>
<td>3.08</td>
</tr>
<tr>
<td>D-transposition of the great arteries</td>
<td>2.40</td>
<td>2.00</td>
<td>1.90</td>
<td>0.00</td>
<td>0.22</td>
<td>6.21</td>
</tr>
<tr>
<td>Double outlet right ventricle</td>
<td>2.20</td>
<td>2.00</td>
<td>2.00</td>
<td>0.00</td>
<td>0.47</td>
<td>6.10</td>
</tr>
<tr>
<td>Truncus arteriosus</td>
<td>0.60</td>
<td>NM</td>
<td>1.40</td>
<td>1.93</td>
<td>5.37</td>
<td>3.63</td>
</tr>
<tr>
<td>Atrioventricular septal defect</td>
<td>4.00</td>
<td>5.40</td>
<td>1.40</td>
<td>1.93</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td>With Down syndrome</td>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Without Down syndrome</td>
<td>1.00</td>
<td>NM</td>
<td>1.93</td>
<td>1.93</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Total anomalous pulmonary venous return</td>
<td>0.60</td>
<td>0.80</td>
<td>NM</td>
<td>0.00</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Ebstein anomaly</td>
<td>0.60</td>
<td>0.50</td>
<td>0.00</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Right obstructive defects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tricuspid atresia</td>
<td>0.30</td>
<td>0.50</td>
<td>0.00</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Pulmonary atresia, intact septum</td>
<td>0.60</td>
<td>0.60</td>
<td>0.00</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Pulmonic stenosis, atresia</td>
<td>5.95</td>
<td>7.30</td>
<td>6.50</td>
<td>58.65</td>
<td>28.20</td>
<td></td>
</tr>
<tr>
<td>Peripheral pulmonary stenosis</td>
<td>7.00</td>
<td>NM</td>
<td>16.35</td>
<td>23.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left obstructive defects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypoplastic left heart</td>
<td>2.10</td>
<td>1.80</td>
<td>0.70</td>
<td>0.00</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Coarctation of aorta</td>
<td>3.50</td>
<td>2.40</td>
<td>3.80</td>
<td>0.96</td>
<td>2.90</td>
<td>3.20</td>
</tr>
<tr>
<td>Aortic arch atresia or hypoplasia</td>
<td>0.60</td>
<td>0.70</td>
<td>0.90</td>
<td>0.00</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Aortic valve stenosis</td>
<td>0.80</td>
<td>2.20</td>
<td>1.50</td>
<td>2.88</td>
<td>16.00</td>
<td></td>
</tr>
<tr>
<td>Mitral insufficiency</td>
<td>1.20</td>
<td>1.20</td>
<td>4.81</td>
<td>35.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Septal defects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VSD</td>
<td>24.98</td>
<td>39.00</td>
<td>45.70</td>
<td>6.73</td>
<td>1.60</td>
<td></td>
</tr>
<tr>
<td>ASD</td>
<td>10.00</td>
<td>7.50</td>
<td>12.20</td>
<td>26.92</td>
<td>2.40</td>
<td></td>
</tr>
<tr>
<td>Isolated hypertrophy</td>
<td>5.00</td>
<td>3.60</td>
<td>11.50</td>
<td>0.00</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Patent ductus arteriosus</td>
<td>8.10</td>
<td>3.80</td>
<td>1.50</td>
<td>5.95</td>
<td>10 of 168; P &lt; .01)</td>
<td></td>
</tr>
<tr>
<td>Other major heart defects</td>
<td>9.70</td>
<td>NM</td>
<td>0.00</td>
<td>0.00</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>90.20</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
<td>NA</td>
<td></td>
</tr>
</tbody>
</table>

The data are listed with the available data on prevalence of cardiac defect in the general population from 3 large studies. Heart defects are grouped following Van Praagh’s system. Eurocat Database Prevalence period spanning from 1999 to 2003: cardiac defects encompass anomalies of cardiac chambers and connections, atrioventricular septal defect, transposition of great vessels, tetralogy of Fallot, malformations of valves, malformations of the great arteries and veins, malformations of cardiac septa, coarctation of aorta, common arterial truncus (www.eurocat.ulster.ac.uk). L-TGA indicates L transposition of the great arteries; NA, not applicable; NM, not mentioned.
of CHD. Mutations in exon 13 account for 18.30%, 55% (10 of 19) being PVS. In Table 4, we compared the frequencies of CHD in patients with NS with their frequencies in the general population, based on registries of congenital anomalies, and computed relative risks (when possible). This relative risk varies widely, from 0.22 for D-transposition of the great arteries to 35.6 for mitral insufficiency.

The spectrum of CHD that is observed with \textit{PTPN11} mutations points to a crucial role for this gene in the normal pattern of right cardiac development and anatomy. Left-sided heart malformations are more prevalent in \textit{PTPN11}-negative patients.

The function of \textit{PTPN11} gene during heart embryogenesis remains to be elucidated. \textit{Ptpn11}, the mouse ortholog of the human \textit{PTPN11}, is required during early mouse development for gastrulation. A mouse that carries the D61G mutation has been produced and recapitulates NS. In a mouse model, Chen et al.\textsuperscript{27} found that normal epithelial growth factor receptor (Egfr) signaling pathway necessitates interaction of Egfr with Shp2. The Egfr pathway is necessary to permit semilunar valve development. Despite their involvement in NS, expression of \textit{PTPN11} does not seem to occur in atrioventricular valves.\textsuperscript{14,27} Mice that bear homozygous \textit{Egfr} mutations show electrocardiographic anomalies, aortic stenosis, and regurgitation. In the chick embryo, the hyperproliferative effect of an expressing Shp2 bearing a type 1 NS mutation Q79R on mesenchymal cells in valve primordia was shown to be mediated by extracellular signal–regulated kinase 1/2 activation through the RAS/MAPK pathway. The effect was similarly obtained when transfecting the gene for a constitutively active MEK-1, a downstream molecule that regulates ERK (extracellular signal–regulated kinase phosphatase. On activation by Ca\textsuperscript{2+}, calcineurin dephosphorylates the NFAT transcription factor, leading to its nuclear translocation and activation. In mice, NFAT acetylates the NFAT transcription factor, leading to its transcriptional activity of NFAT, pointing to a possible link between the calcium/calcineurin/NFAT pathway and \textit{PTPN11}.\textsuperscript{30} It is interesting that \textit{PTPN11} is not involved in nonsyndromic atrioventricular septal defects and coarctation of the aorta\textsuperscript{34} or in nonsyndromic HCM.\textsuperscript{35}

\textbf{CONCLUSIONS}

Our study of CHDs in type 1 NS illustrates the wide spectrum of developmental anomalies that result from dysregulation of the RAS/MAPK signaling pathway and the difficulty of correlating molecular anomalies with the occurrence of a type of CHD.

\textbf{REFERENCES}

20. Kontaridis MI, Swanson KD, David FS, Barford D, Neel BG. PTPN11 (Shp2) mutations in LEOPARD syndrome have dominant negative, not activating, effects. J Biol Chem. 2006;281:6785–6792
Families’ Health-Related Social Problems and Missed Referral Opportunities

Eric W. Fleegler, MD, MPH\textsuperscript{a,b}, Tracy A. Lieu, MD, MPH\textsuperscript{b,c,d}, Paul H. Wise, MD, MPH\textsuperscript{e}, Sharon Muret-Wagstaff, PhD, MPAb,d

\textsuperscript{a}Division of Emergency Medicine and \textsuperscript{b}Department of Medicine, Children’s Hospital Boston, Boston, Massachusetts; \textsuperscript{c}Harvard Medical School, Boston, Massachusetts; \textsuperscript{d}Center for Child Health Care Studies, Department of Ambulatory Care and Prevention, Harvard Pilgrim Health Care, Boston, Massachusetts; \textsuperscript{e}Department of Pediatrics, Stanford University School of Medicine, Stanford, California

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

ABSTRACT

OBJECTIVES. The objectives of this study were to characterize (1) families’ cumulative burden of health-related social problems regarding access to health care, housing, food security, income security, and intimate partner violence; (2) families’ experiences regarding screening and referral for social problems; and (3) parental acceptability of screening and referral.

METHODS. We surveyed 205 parents of children who were 0 to 6 years of age and attended 2 urban pediatric clinics for a well-child visit using a self-administered, computer-based questionnaire. The questionnaire included previously validated questions about health-related social problems and new questions about screening and referral in the past 12 months.

RESULTS. A total of 205 (79%) of 260 eligible families participated. Eighty-two percent of families reported ≥1 health-related social problem; 54% experienced problems in ≥2 social domains. Families experienced similar types and frequencies of problems despite demographic differences between clinics. One third of families reported no screening in any domain in the previous 12 months. Of 205 families, 143 (70%) identified at least 1 need for a referral; 101 (49%) expressed ≥1 unmet referral need. Of families who reported receiving referrals, 115 referrals were received by 79 families; of the referrals made, 63% (73 of 115) led to contact with the referral agency, and 82% (60 of 73) of the referral agencies were considered helpful. A computer-based system in a pediatrician’s office for future screening and referral for health-related social problems was deemed acceptable by 92% of parents.

CONCLUSIONS. Urban children and families reported a significant burden of health-related social problems yet infrequent pediatric screening or referral for these problems. Of families who reported receiving referrals, a majority contacted the recommended agencies and found them helpful. This study also demonstrates the feasibility of using a computer-based questionnaire to identify health-related social problems in a routine outpatient clinic setting.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-1505
doi:10.1542/peds.2006-1505

Results from this study were presented at the annual research meeting of AcademyHealth, June 7, 2004, San Diego, CA; and the annual meeting of the Pediatric Academic Societies, May 16, 2005, Washington, DC.

Key Words
social problems, screening, referral, access to health care, housing, hunger, income, domestic violence, computer, pediatric

Abbreviations
AAP—American Academy of Pediatrics
HRSP—health-related social problem
AHP—academic health practice
CHC—community health center
CEHP—Childhood Community Hunger Identification Project
FPL—federal poverty level
CI—confidence interval

Accepted for publication Nov 27, 2006

Address correspondence to Eric W. Fleegler, MD, MPH, Division of Emergency Medicine, MA-001, Children’s Hospital Boston, 300 Longwood Ave, Boston, MA 02115. E-mail: eric.fleegler@childrens.harvard.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics

e1332 FLEEGLER et al
The American Academy of Pediatrics (AAP) states that its mission “is to attain optimal physical, mental and social health and well-being for all infants, children, adolescents and young adults.” To fulfill this charge, pediatricians must address medical problems as well as social and economic issues that can adversely affect health. Children are particularly vulnerable to problems such as food insecurity and substandard housing, and their effects can be lifelong. Providing assistance for families’ health-related social problems (HRSPs) requires systematic screening and referral.

Although the AAP recommends screening for major social problems in primary care settings, the extent of comprehensive screening for HRSPs in routine practice is unknown. The rate of screening for intimate partner violence alone may be as low as 5%. Whereas parents’ approval of screening for intimate partner violence is well documented, no research published to date has evaluated parents’ support for comprehensive screening for HRSPs in pediatric settings.

In this study, families with young children provided self-assessments of HRSPs by completing a computer-based questionnaire before a clinic visit. We specifically evaluated (1) the presence of HRSPs in 5 social domains, (2) the frequency of pediatric screening and referral for these problems, and (3) families’ attitudes toward screening and referral for HRSPs.

METHODS

Design and Participant Selection
This is a cross-sectional, descriptive study of families’ self-assessed health-related social problems and referral needs conducted in August to September 2003 at 2 outpatient pediatric clinics in Boston: 1 academic health practice (AHP) and 1 community health center (CHC). Each family was represented by a primary caregiver. Adults in clinic waiting areas were screened consecutively for HRSPs by completing a computer-based questionnaire before a clinic visit. We specifically evaluated (1) the presence of HRSPs in 5 social domains, (2) the frequency of pediatric screening and referral for these problems, and (3) families’ attitudes toward screening and referral for HRSPs.

Survey Instrument and Measures
The survey instrument focused on HRSPs and experiences and opinions related to screening and referral for social problems (Table 1; the original questionnaire may be viewed at www.onlineadvocate.org). A possible inclusion list of 25 social domains was initially derived using literature review and key informant interviews with health and social services experts. A modified Delphi technique was used to select the top 5 most relevant topics for inclusion. The 5 health-related social domains were (1) access to health care, (2) housing, (3) food security, (4) income security, and (5) intimate partner violence. Confidentiality and anonymity were protected and emphasized throughout the computer survey pro-

### Table 1: Survey-Item Categories

<table>
<thead>
<tr>
<th>Health-Related Social Domains</th>
<th>Previous Experiences</th>
<th>Parental Opinions</th>
<th>Demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access to health care</td>
<td>Social problems</td>
<td>Acceptability of screening individual social domains</td>
<td>Parent and child ages</td>
</tr>
<tr>
<td>No health insurance</td>
<td>Screened</td>
<td>Acceptability of using computer to screen and refer in pediatric office</td>
<td>Parent and child genders</td>
</tr>
<tr>
<td>Missed medical care</td>
<td>Referred</td>
<td>Comfort using computer</td>
<td>Education level</td>
</tr>
<tr>
<td>Missed prescriptions</td>
<td>Referral agencies</td>
<td>Distance willing to travel</td>
<td>Race/ethnicity</td>
</tr>
<tr>
<td>Housing</td>
<td>Followed up</td>
<td></td>
<td>Marital status</td>
</tr>
<tr>
<td>Homeless or doubled up</td>
<td>Helpfulness</td>
<td></td>
<td>Immigration status</td>
</tr>
<tr>
<td>Utilities shut off</td>
<td></td>
<td></td>
<td>Household status</td>
</tr>
<tr>
<td>Major housing problem</td>
<td></td>
<td></td>
<td>Transportation composition</td>
</tr>
<tr>
<td>Food security</td>
<td></td>
<td></td>
<td>Income</td>
</tr>
<tr>
<td>Food-insecure or hungry</td>
<td></td>
<td></td>
<td>Language</td>
</tr>
<tr>
<td>Income security</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed and looking for work</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intimate partner violence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal abuse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical abuse</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PEDIATRICS Volume 119, Number 6, June 2007 e1333
cess. The survey did not collect any identifying data or contain any links to the informed consent.

The survey instrument combined previously validated questions and scales to assess the social domains as well as new questions to assess experience with social problem screening and referral. Questions in the health-related social domains evaluated the presence and the extent of problems, use of available social services, and barriers to access that families experienced. The survey used a response-driven, branched questionnaire that ranged from 90 to 166 questions.

The access to health care domain included questions from the National Health Interview Survey,13 Behavioral Risk Factor Surveillance System,14 and Child Care Experience and Needs Questions15 and assessed for both the parent and the child their health insurance status (presence/absence and type of health care coverage, reason without health insurance), use of primary care providers and usual source of care, and problems receiving medical services and/or medications within the previous 12 months (and reason for problem). An access to health care problem included either current lack of health insurance or inability to receive medical care or fill a prescription during the previous 12 months for the parent or the child.

The housing domain included questions from the American Housing Survey16 and assessed household size and makeup, current housing status (own/rent/doubled up/homeless), cost of housing, concerns about impending eviction, and previous 12-month experiences with homelessness/doubled up (presence and duration), housing utilities (threatened/shut off/receipt of fuel assistance), and housing hazards. Housing hazards within the past 12 months included: roof leaked, problems with electrical wiring, no heat for >24 hours, and water leaks in the home from inside (pipes, sinks, toilets) and outside (walls, roof). Housing hazards within the past 3 months included: none of the toilets worked, any rats or mice in home or building, any cockroach/insect infestation, no running water in the house, and broken utilities. A housing problem included either currently homeless or doubled up, utilities shut off during the previous 12 months, or a major structural housing problem as defined by the American Housing Survey and validated in previous research studies.17

The food security domain included the previously validated 8-point food security scale from the Childhood Hunger Identification Project (CHIP).18–20 The CHIP asks the parent to think about the past 12 months and answer questions about (1) running out of money to buy food, (2) using a limited number of foods to feed the family, (3) adults eating less than they should, (4) adults skipping meals, (5) children eating less than they should, (6) children saying that they were hungry because of lack of food, (7) children skipping meals, and (8) children going to bed hungry because of lack of food. Additional questions assessed use of food stamps and participation in the Supplemental Nutrition Program for Women, Infants, and Children. A food security problem was defined as food insecure (1–4 positive responses) or frank hunger (5–8 positive responses) according to the CHIP scale.

The income security domain used questions from the Philadelphia Survey of Work and Family21 and surveys noted previously.13–15 Questions assessed employment status (employed/self-employed/out of work [duration]/homemaker/student/retired/unable to work), reason for difficulty finding/maintaining work, use of job training/job placement/interest in getting a paid job, and household income, as well as use, duration, and amount of welfare.22 Supplemental Security Income, and child support. Additional questions assessed problems with and reasons for missing work and for missing medical appointments. An income security problem was narrowly defined as being currently unemployed and looking for work.

The intimate partner violence domain included an introduction and 3 questions that were used previously to screen women with young children.14 These questions included the following: “In the past year, have you been emotionally or verbally abused by your partner or someone important to you; for example, has anyone sworn at you, threatened you, or threatened to throw something at you?” “In the past year, have you been hit, slapped, kicked, or otherwise physically hurt by your partner or someone close to you?” and, “Do you feel safe in your current relationship?” Additional questions assessed the type of intimate partner violence experienced and use of medical care or services related to violence. An intimate partner violence problem was defined as verbal or physical abuse during the previous year.

Previous experience with screening and referral was assessed with a series of up to 5 questions for each social domain. Referral need in each domain was ascertained using the following question: “In the past 12 months, have you been given a referral to an agency to help you with your [domain category]?” A response of “yes” or “no, but I wanted a referral,” defined a referral need (versus, “No, and I did not want a referral” or “not sure”). A question at the end of the survey asked, “Think about the survey you have just completed. How would you feel about taking a computer-based survey in your doctor’s waiting room that evaluated a family’s social issues and made referrals to local agencies?” Reply options were (1) welcome it, (2) not mind at all, (3) be mildly annoyed, and (4) be very annoyed.10

**Computer Program**

The computer program was adapted from the Promote Health Survey11,12 with permission of the principal investigator. The forced-answer, branched questionnaire asked questions that were personalized with the child’s
first name using a single question per screen. The survey consisted of yes/no, multiple-choice, fill-in-the-blank, and checklist questions. A figure on the bottom of the screen tracked the percentage of completion of the survey. Data were collected and stored using Microsoft Access and Excel 2002/SP-2 (Microsoft Corp, Redmond, WA).

Translation and Testing
The survey wording was translated, back-translated, and refined by professional translators who represented 6 different Spanish-speaking countries. The English version was at a fifth- to sixth-grade reading level and the Spanish version was at a fourth- to fifth-grade reading level.23 A focus group of 13 volunteer parents and health care providers pilot-tested the computerized questionnaire for usability, content, and construct validity.

Analysis
The primary outcome measures were the percentages of families who (1) experienced HRSPs; (2) received screening for social problems, received needed referrals, and found referral agencies helpful; and (3) reported willingness to use computer-based screening and referral for social problems in pediatric settings in the future. Although demographics differed between the 2 clinic sites, the frequencies of HRSPs, the experience with referrals, and acceptability of screening were similar. Therefore, data from the 2 clinic settings were combined to provide a descriptive analysis of the families’ burden of HRSPs, parental experiences with screening and referral, and parental opinions. Between-site differences in demographics and all outcome measures were tested by using the \( \chi^2 \) and Student’s \( t \) test when appropriate and are noted when statistically significant differences exist.

Differences in presence of social problems by demographics and household characteristics were tested by using the \( \chi^2 \) test. All statistical tests were 2-tailed and considered significant at \( P < .05 \). All analyses were performed using SPSS 11.1 (SPSS, Chicago, IL).

RESULTS

Study Population
During a 5-week period, trained research assistants approached 450 families. A total of 190 did not meet inclusion criteria. Among the 260 eligible parents, 79% (205 of 260) agreed to participate. Ninety-four percent (193 of 205) of participants completed the entire survey, and an additional 5 completed the HRSP questions but not the demographics (Fig 1). Demographics of eligible parents who refused are not available.

Thirty-four percent (70 of 205) of parents took the survey in Spanish. Of the 193 surveys completed, average time to completion was 20 minutes, and participants answered an average of 123 questions (interquartile range: 117–128). Table 2 depicts demographics of the surveyed participants. Overall, 57% of parents were Hispanic and 29% were black. Among respondents, 34% had not completed high school and 62% were immigrants; 62% had a family income at or below the poverty level.

HRSPs
Of the 198 families who completed all 5 social domains, 162 (82%) had ≥1 HRSP. Comorbidity was high: of the 198 families, 28% experienced 1 HRSP, 32% experienced 2 HRSPs, and 22% experienced ≥3 HRSPs. Demographic variables that correlated with higher risk for having an HRSP are shown in Table 2 and included

![Figure 1](#)
Forty-five percent of families (95% confidence interval [CI]: 38–52) had a problem with access to health care, including no health insurance, unable to receive medical care, and unable to fill a prescription. In addition, 22% relied on “routine care” from either an emergency department or urgent care center or lacked a place for usual medical care.

Housing problems were the most prevalent HRSP (56%; 95% CI: 48–62), with a higher prevalence among the CHC participants (62%) compared with the AHP participants (48%; P = .034). Fifteen percent of the families were either homeless or doubled up; an additional 10% had experienced homelessness within the past year, and 22% expressed concern about being evicted. Eight percent of families had had their utilities shut off during the previous 12 months, and an additional 10% had had their utilities threatened. Forty-four percent of families had had their utilities shut off during the previous 12 months, and an additional 10% had experienced homelessness within the past year, and 22% expressed concern about being evicted. Eight percent of families had had their utilities shut off during the previous 12 months, and an additional 10% had experienced homelessness within the past year, and 22% expressed concern about being evicted. Eight percent of families had had their utilities shut off during the previous 12 months, and an additional 10% had experienced homelessness within the past year, and 22% expressed concern about being evicted. Eight percent of families had had their utilities shut off during the previous 12 months, and an additional 10% had experienced homelessness within the past year, and 22% expressed concern about being evicted.

Table 3 depicts the specific HRSPs experienced by the families. Problem prevalence in each social domain was based on the number of families who completed that domain’s questions. Statistically significant differences between the sites for specific social problem prevalence occurred in the housing and income domains, as noted next.

Forty-five percent of families (95% confidence interval [CI]: 38–52) had a problem with access to health care, including no health insurance, unable to receive medical care, and unable to fill a prescription. In addition, 22% relied on “routine care” from either an emergency department or urgent care center or lacked a place for usual medical care.

### Table 2: HRSPs Experienced by Families (n = 193) in the Preceding 12 Months

<table>
<thead>
<tr>
<th>Demographic</th>
<th>Total (n = 193), n (%)</th>
<th>≥1 HRSP, n (%)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (SD) age of parent, y</td>
<td>29.3 (6.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD) age of child, y</td>
<td>21.1 (1.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parent gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>174 (90)</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>19 (10)</td>
<td>79</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Child gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>100 (52)</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>93 (48)</td>
<td>81</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Parent race/ethnicity*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>110 (57)</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>55 (29)</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>White/mixed/other</td>
<td>28 (15)</td>
<td>64</td>
<td>.043</td>
</tr>
<tr>
<td>Parent education*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No high school degree</td>
<td>66 (34)</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>High school diploma or GED</td>
<td>58 (30)</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td>Some college or more</td>
<td>70 (36)</td>
<td>73</td>
<td>.066</td>
</tr>
<tr>
<td>Parent immigration status*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undocumented/unsure</td>
<td>28 (15)</td>
<td>93</td>
<td></td>
</tr>
<tr>
<td>Legal immigrant</td>
<td>90 (47)</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td>Born in US</td>
<td>75 (39)</td>
<td>75</td>
<td>.087</td>
</tr>
<tr>
<td>Parent relationship status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single, no partner</td>
<td>61 (32)</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>Single, living separate from partner</td>
<td>31 (16)</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>Single, living with partner</td>
<td>30 (16)</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>71 (37)</td>
<td>72</td>
<td>.021</td>
</tr>
<tr>
<td>Median household size, No. of people</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income, median, $5</td>
<td>15,314</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No income reported</td>
<td>24 (12)</td>
<td>96</td>
<td></td>
</tr>
<tr>
<td>≤15,000</td>
<td>84 (44)</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>&gt;15,000</td>
<td>85 (44)</td>
<td>72</td>
<td>.006</td>
</tr>
<tr>
<td>FPL*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not calculable*</td>
<td>31 (16)</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>≤50 %</td>
<td>53 (28)</td>
<td>87</td>
<td></td>
</tr>
<tr>
<td>≤100 %</td>
<td>37 (19)</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>≤200 %</td>
<td>36 (19)</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>&gt;200 %</td>
<td>36 (19)</td>
<td>52</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

*GED indicates general equivalency diploma.

a Statistically significant difference between clinics (P < .05) by χ² or Student’s t test.

b No income or household size was reported.

c Food security problem according to clinic: 33% (AHP) and 43% (CHC), P < .05 by Student’s t test.

d Housing problems according to clinic: 48% (AHP) and 62% (CHC), P < .05 by Student’s t test.

### Table 3: Prevalence of HRSPs

<table>
<thead>
<tr>
<th>Parameter</th>
<th>n (%)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of HRSPs (N = 198)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>36 (18)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>55 (28)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>63 (32)</td>
<td></td>
</tr>
<tr>
<td>≥3</td>
<td>44 (22)</td>
<td></td>
</tr>
<tr>
<td>Domain and problem (N = 198–203)*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Access to health care (N = 203)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No health insurance for parent</td>
<td>43 (21)</td>
<td>13–26</td>
</tr>
<tr>
<td>No health insurance for child</td>
<td>19 (9)</td>
<td>5–13</td>
</tr>
<tr>
<td>Unable to receive medical care</td>
<td>39 (19)</td>
<td>13–24</td>
</tr>
<tr>
<td>Unable to fill prescription</td>
<td>37 (18)</td>
<td>12–23</td>
</tr>
<tr>
<td>Housing (N = 201)*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homeless or living in homeless shelter</td>
<td>14 (7)</td>
<td>3–10</td>
</tr>
<tr>
<td>Doubled up</td>
<td>16 (8)</td>
<td>4–11</td>
</tr>
<tr>
<td>Utilities shut off</td>
<td>17 (8)</td>
<td>4–12</td>
</tr>
<tr>
<td>Major structural housing problems (≥1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>42 (21)</td>
<td>15–26</td>
</tr>
<tr>
<td>2</td>
<td>27 (13)</td>
<td>8–18</td>
</tr>
<tr>
<td>≥3</td>
<td>19 (9)</td>
<td>5–13</td>
</tr>
<tr>
<td>Food security (N = 201)*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCHIP 1–4 (food-insecure)</td>
<td>56 (28)</td>
<td>21–34</td>
</tr>
<tr>
<td>CCHIP 5–8 (hungry)</td>
<td>22 (11)</td>
<td>6–15</td>
</tr>
<tr>
<td>Income security (N = 198)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed and looking for work</td>
<td>34 (17)</td>
<td>11–22</td>
</tr>
<tr>
<td>Intimate partner violence (N = 198)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal violence</td>
<td>28 (14)</td>
<td>9–19</td>
</tr>
<tr>
<td>Physical violence</td>
<td>26 (13)</td>
<td>8–17</td>
</tr>
<tr>
<td>Physical violence within past month</td>
<td>13 (7)</td>
<td>3–10</td>
</tr>
<tr>
<td>Physical violence within past month</td>
<td>7 (4)</td>
<td>0–6</td>
</tr>
</tbody>
</table>

* Domain number was based on completion of specific survey domain.

a Housing problems according to clinic: 48% (AHP) and 62% (CHC), P < .05 by Student’s t test.

b No income or household size was reported.
(score 5–8; limited food affecting the parent and the child). Eleven percent of families stated that they did not have “enough food to eat today.” Food subsidy use was high; 54% of families were enrolled in the Supplemental Nutrition Program for Women, Infants, and Children, and 40% received food stamps.

Income insecurity was narrowly defined as being “unemployed and looking for work”; 17% of families (95% CI: 11–22) met this criterion. An additional 34 (17%) parents stated that they were interested in getting a paid job. A total of 108 (55%) families had an income of $15,000 or less or no reported income (61% ≤100% FPL). Twelve percent of parents reported that their child did not go to the doctor because the parent could not leave school or work, and 18% of parents reported that their child had not gone to the doctor because of financial concerns.

Fourteen percent of parents (95% CI: 9–19) had experienced intimate partner violence within the past year; 13% reported being “emotionally or verbally abused,” including threats of harm, and 7% reported physical violence. Four percent had experienced physical violence within the past month. Seven percent of parents reported not feeling safe in their current relationship. Of the 28 parents who experienced intimate partner violence, only 5 (18%) had seen a doctor secondary to abuse.

**Screening, Referral Need, and Referral Agency Experience**

Parent-reported screening rates in the previous year were low in all domains. Screening rates by health-related social domains were: access to health care, 23%; housing, 31%; food security, 17%; income security, 21%; and intimate partner violence, 36%. Comprehensive screening of families was minimal. Of the 198 families who completed all 5 domains, only 2% (5 of 198) had been screened in all 5 domains, and 32% (66 of 198) had not been screened in any domain. Table 4 shows the screening rate, referral need, referral receipt, and referral agency experience. Screening rates between clinics differed only in the housing domain (22% [AHP] vs 39% [CHC]; P = .037).

Referral need (referral received or wanted) ranged from a low of 15% for intimate partner violence to a high of 44% for housing. For each domain, referral receipt percentage was calculated with respect to the total number of referrals needed. In aggregate, 143 (70%) of 205 families identified a need for referral in ≥1 domain, and 101 (49%) of 205 had ≥1 unmet referral need. With the exception of access to health care, fewer than half of families received needed referrals in each domain, with a low of 14% for income security and a high of 65% for access to health care.

Of families who reported receiving referrals, the majority contacted the referral agency; the lowest contact rate was for income referrals (55%), and the highest contact rate was for housing referrals (68%). The majority of referred parents found the referral agencies helpful, with the 1 notable exception of agencies that assist with income security, which were found helpful by only 17% (1 of 6) of families. Families found other referral agencies significantly more helpful: helpfulness of agencies that assist with access to health care, 92% (24 of 26); with housing, 71% (12 of 17); with food security, 94% (15 of 16); and with intimate partner violence, 100% (8 of 8).

**Screening Desirability**

More than 80% of parents said that they would “welcome” or “not mind at all” inquiries about problems within each domain (Fig 2). When asked specifically about the acceptability of using a computer system to screen and refer families for HRSPs at the pediatrician’s office during a well-child visit, 92% responded that they would “welcome it” or “not mind at all.”

**DISCUSSION**

This study found that despite high rates of HRSPs, reported screening and referral rates were low among families in 2 urban populations. Our research was unique in that it demonstrated the feasibility of using a computer-based questionnaire in pediatric outpatient clinical settings to ask families about HRSPs and referral needs.

**Prevalence**

Within each social domain, the problems that were identified in this population reflect similar findings in previ-

---

**TABLE 4** HRSP Screening, Referral Need/Receipt, and Referral Agency Experience Within Past 12 Months

<table>
<thead>
<tr>
<th>Domain</th>
<th>Respondents, N</th>
<th>Screened, n (%)</th>
<th>Needed, n (%)a</th>
<th>Received, n (%)b</th>
<th>Contacted, n (%)b</th>
<th>Helpful, n (%)b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access to health care</td>
<td>203</td>
<td>47 (23)</td>
<td>63 (31)</td>
<td>41 (65)</td>
<td>26 (63)</td>
<td>24 (92)</td>
</tr>
<tr>
<td>Housing</td>
<td>202</td>
<td>63 (31)c</td>
<td>89 (44)</td>
<td>25 (28)</td>
<td>17 (68)</td>
<td>12 (71)</td>
</tr>
<tr>
<td>Food security</td>
<td>201</td>
<td>35 (17)</td>
<td>68 (34)</td>
<td>24 (35)</td>
<td>16 (67)</td>
<td>15 (94)</td>
</tr>
<tr>
<td>Income security</td>
<td>198</td>
<td>42 (21)</td>
<td>77 (39)</td>
<td>11 (14)</td>
<td>6 (55)</td>
<td>1 (17)</td>
</tr>
<tr>
<td>Intimate partner violence</td>
<td>198</td>
<td>71 (36)</td>
<td>30 (15)</td>
<td>14 (47)</td>
<td>8 (57)</td>
<td>8 (100)</td>
</tr>
</tbody>
</table>

a Needed indicates referral received or wanted.

b Percentage of previous column.

c Screening for housing problem according to clinic: AHP 22% and CHC 39%, P < .05 by χ².
ous studies.\textsuperscript{24-29} It is notable that 45% of families have a problem with access to health care, although they were screened within a health care setting. This is likely a reflection of a high level of access afforded children via MassHealth (Medicaid) and the State Children’s Health Insurance Program but denied to the parents who lack health insurance. Although the correlation between low income and high rates of HRSPs is not surprising, the frequent overlap of multiple problems is significant and likely compounds their impact on pediatric families. Within this study, 54% of families experienced $\geq 2$ major HRSPs, an important finding that was elucidated by comprehensive rather than single-issue screening. Populations that are screened in nonmedical settings may have even higher rates of HRSPs, because families who are screened within a pediatric clinic already have some level of demonstrated access to health care.

Identification and Referral for HRSPs
Pediatricians may underestimate the prevalence of HRSPs and the value of providing referrals. Our finding that 49% of families have an unmet referral need demonstrates a gap between what families need and how their physicians are responding. National and local studies have demonstrated screening and referral rates for intimate partner violence as low as 5%\textsuperscript{5,10}; no other known studies have evaluated screening for other major social problems.

Within this study, when families received referrals, 63% (73 of 115) contacted the agency, and of those, 82% (60 of 73) of the agencies contacted were considered helpful. Previous studies in the medical, social work, and psychology literature support the utility of referral interventions, including increased rates of health insurance,\textsuperscript{31,32} improved food security, and increased economic resources.\textsuperscript{33}

Feasibility of Computer-Based Screening and Referral
Can a computer-based screening and referral system within a pediatrician’s office connect families to referral agencies? Sixty-two percent of families stated that they welcomed such a tool, and an additional 30% said that they would not mind at all using it. The Health Belief Model, which informed the conceptual development of
this project, describes the probability of an individual taking action to get help with a health problem as a balance among (1) an individual’s perception of his or her susceptibility to and seriousness of the problem, (2) modifying factors that include perceived threats and cues to actions, and (3) the perceived benefits minus the perceived barriers to taking action. A computer-based system that assesses HRSPs and provides feedback as well as actual referrals has the capacity to reinforce a family’s awareness of the HRSPs, provide cues to action that can be reinforced by the pediatrician, explain the benefits of following up with the agencies, and, finally, reduce barriers by providing printed referrals that help families connect with appropriate agencies. Future research will evaluate the use and utility of a fully integrated screening and referral system and will assess families’ desire to share the provided information with their physician and whether to include it within the medical chart.

Previous studies of computer use in medical settings support its acceptance. A study of computer use by families in medical settings demonstrated >90% support for use of a computer before each medical visit to improve health services delivery. Other studies have shown that patient-centered computer systems are an effective means to obtain medical histories, to make clinical decisions, to develop asthma action plans, to improve parental knowledge of key pediatric issues, and to improve overall delivery of pediatric primary care. As stated, this study was performed within the waiting rooms of an AHP and a CHC. Although barriers to using a computer in these settings, including crowded conditions and minimal privacy, may seem extremely challenging, 79% of eligible families participated and 94% of participating families completed the survey. We believe that this speaks to the strong desire of families to use computer systems that are designed to help them and their faith in privacy that is afforded to them by the privacy screens and the reinforced confidential nature of the tool. Of note, during the study, both doctors and nurses thanked the research assistants for giving their patients something meaningful to do while waiting to be seen.

Strengths and Limitations

The primary strength of this study was its comprehensive analysis of HRSPs in a broad array of domains collectively. In addition, it was conducted in the real-world settings of community and academic-based clinics using a computer. Although the study is limited by a relatively small sample size from 2 clinics in a single city, the populations studied reflect 2 of the largest minority ethnic groups in the United States (black and Hispanic). Both of these clinics have onsite social workers and missions to serve the underserved; therefore, the extent of social problems and referral rates may actually be greater than elsewhere. The generalizability of these findings to other clinical practices in other cities with potentially different social problems and a different range of social service agencies available will need additional study.

Self-reported responses may be affected by social desirability bias. The effect of social desirability bias might have led to families’ minimizing problems or giving higher ratings than warranted on helpfulness of social services. However, previous studies of computer-based evaluation of sensitive and personal issues suggested that high rates of “honest” reporting can be expected by this modality. Recall bias may lead parents to underestimate (or overestimate) whether they had been screened by a clinician for each of the HRSPs.

Another limitation is the lack of formal assessment of the burden on the clinics of screening families. However, active family self-assessment using a computer in the waiting room is likely to be more systematic and less time-consuming for the clinician than traditional assessment by interview.

CONCLUSIONS

Urban families with young children bear a significant burden of HRSPs that remain largely unobserved and unattended by pediatric practices. Among families who reported receiving referrals, most said that they made contact with designated agencies and found the agencies helpful in addressing these challenges. Pediatric practices have the potential to play an important role in enabling families to identify HRSPs and receive referrals to community resources. Finally, this study demonstrates the feasibility of using a computer-based system to implement family self-assessment in a routine outpatient setting and may thereby help address some of the barriers in providing these assessments and referrals in pediatric practices.

ACKNOWLEDGMENTS

Dr Fleegler was supported by Agency for Healthcare Research and Quality grant T32 HS000063 to the Harvard Pediatric Health Services Research Fellowship Program. We are grateful for additional funding support provided by the American Medical Association Foundation Seed Grant, Boston Medical Center Department of Pediatrics Kids’ Fund, and Children’s Hospital Boston, Division of General Pediatrics.

Special thanks go to Karin Rhodes, MD, MPH, the principal investigator who developed the “Promote Health Survey,” for providing access to the original program and ongoing support and encouragement. We are grateful to Derek Roth Gordon for computer program-
REFERENCES


Visual Function at 11 Years of Age in Preterm-Born Children With and Without Fetal Brain Sparing

Joke H. Kok, PhD, MDa, Liesbeth Prick, PhD, MDb, Elly Merckelb, Yolande Everhardb, Gijs J. Q. Verkerkc, Sicco A. Scherjon, PhD, MDd

Departments of aNeonatology, bOphthalmology, and cRehabilitation, Emma Children’s Hospital and Academic Medical Center, Amsterdam, Netherlands; dDepartment of Obstetrics, Leiden University Medical Center, Leiden, Netherlands

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

ABSTRACT

OBJECTIVE. We have demonstrated earlier an accelerated maturation of the visual evoked potential in the first year of life in preterm infants with antenatal brain sparing. We have now assessed visual functioning at 11 years of age in the same cohort and compared the groups with and without brain sparing.

DESIGN/METHODS. One hundred sixteen survivors included in a study on the outcome of preterm infants born at <33 weeks’ gestation with and without fetal brain sparing and admitted to the NICU were followed extensively. Ninety-eight infants (85%) were again assessed at 11 years of age. Data were available for fetal Doppler measurements indicating brain sparing, neonatal cerebral ultrasound scanning, and developmental outcome in the first 5 years. Mean birth weight was 1303 g; mean gestational age was 29.8 weeks. The infants were divided into 2 groups with and without brain sparing. Visual functioning was estimated by measuring visual acuity, visual fields, eye position, and binocular function and by visual motor tests.

RESULTS. Six percent of the children were found to have a visual acuity of <0.8, 12% had strabismus, and 14% to 46% showed abnormal results on the visual motor tests. No statistical differences were found between the 2 groups. However, children with severe cerebral ultrasound diagnoses in the neonatal period were found to have significantly more abnormalities on visual functioning and lower scores on visual motor tests than children without these morbidities.

CONCLUSIONS. Children with fetal brain sparing do not demonstrate a different development of their visual functioning at late school age. However, an abnormal cerebral ultrasound in the neonatal period is associated with impaired visual function in later life.
Preterm birth with or without intraterine growth restriction (IUGR) is associated with significant mortality and morbidity both in the neonatal period and in later life. New approaches, both in obstetric and neonatal treatment of the preterm-born infant introduced after 1985, such as the use of antenatal corticoids and surfactant, as well as new modes of respiratory support and regionalization of care, have improved outcome.

Because of placental insufficiency, a reduction of umbilical blood flow is seen in pregnancies with IUGR, whereas the blood flow to the brain is preserved by a compensatory mechanism. This centralizing of blood flow to the fetal brain is called the “brain-sparing effect.” This term may be somewhat misleading. It suggests a relative protection of the brain during fetal development and does not necessarily reassure a normal developmental outcome. Especially in case of decompenstation, the hemodynamic changes are found to be associated with fetal hypoxia and possibly even with adverse neurodevelopmental outcome.

Cranial ultrasound has been shown to be very useful in the prediction of neurodevelopmental outcome. Emphasis has moved to define not only intracranial hemorrhage (ICH) and ventricular dilation but also white matter injury. Neonates with ICH but especially neonates with ischemic white matter lesions are prone to visual-acuity deficits. A much earlier confrontation with visual stimuli during a period of rapid maturation and the possible deleterious effects of illnesses sustained during the perinatal period place the preterm-born infant at significant risk for impairment of the visual system. Besides the widely known retinopathy of prematurity (ROP), the incidence of refractive errors, strabismus, amblyopia, and cortical visual impairment is reported to be high (≥40%) in preterm-born infants, especially when (cystic) periventricular leukomalacia was present. Although severe visual loss is rare in preterm-born infants, little is known of minor impairment of visual function and defects of the visual fields. The quality of the visual input may be important for optimal visual development, as well as for other fields of neurodevelopment, most probably by alterations at the level of neuronal connectivity. Failure of the myelination process, related to defects in the visual tract, might contribute to poor developmental outcome. Myelination of the optic nerve and tract is incomplete at term and continues during the first 2 postnatal years.

Although we have demonstrated that redistribution of the circulation to the brain is associated with IUGR, there was no independent association with neurologic outcome at 3 years of age. This was also confirmed in other studies, in which neurologic outcome at 5 years could not be predicted by antenatal Doppler studies.

We found unexpectedly an accelerated maturation of visual evoked potentials in the first year of life in infants who showed fetal brain sparing as a result of hemodynamic adaptation to placental insufficiency, independent of gestational age at delivery. Also, we found an abnormal visual motor integration (VMI) on the RAKIT (Revision of the Amsterdam Children’s Intelligence Test) at 5 years of age. Because of the essential role of visual function in motor development and cognitive development, this study was designed to investigate whether preterm-born infants with brain sparing in the fetal period are at higher risk to develop visual function disturbances at school age than preterm infants without fetal brain sparing. We also analyzed the influence of other perinatal risk factors on visual outcome.

METHODS

The study group has been described in more detail in previously published studies. The initial cohort of 128 preterm-born infants was formed from infants who were admitted to the NICU of the Academic Medical Center, University of Amsterdam, in 1989 and who were included in a perinatal study on the effects of fetal brain sparing on infant outcome. From March 1989 until December 1989, all mothers with a threatening preterm delivery (>25 and <34 weeks), both singleton and multiple births, were included in the study. A total of 116 infants survived the first year of life and were extensively followed until 5 years of age. Almost all children were assessed at 6, 12, and 24 months and 3 and 5 years of age to evaluate outcome of different developmental domains. For this study, children were invited to visit the outpatient clinic at the age of 11 years. Approval was obtained by the medical ethical review board of our hospital. Informed consent was obtained from all parents.

Clinical Protocol and Previous Follow-up Data

Prenatal Doppler Measurements

The technique of Doppler investigation was described previously. Briefly, measurements from the umbilical artery and the middle cerebral artery were performed within 1 week before delivery. Doppler measurements were not available for attending obstetricians and were not used for timing of delivery. Pulsatility index of the umbilical artery and the middle cerebral artery were recorded, and the ratio between them (the umbilical/cerebral [U/C] ratio) was calculated. We considered a U/C ratio >0.72 as fetal brain sparing. The last Doppler measurement before delivery was used for statistical analysis. Fetuses with chromosomal disorders or major congenital abnormalities were excluded from the study.

Data on Delivery

Birth weight and birth weight ratio, calculated according to the Dutch growth curves and gestational age at delivery, use of antenatal steroids, multiple pregnancy, and Apgar score at 5 minutes were obtained from the original study database.
Neonatal Data
Data on respiratory distress syndrome, surfactant use, oxygen dependence at 28 days after birth, sepsis, and cranial ultrasound diagnosis were obtained from the original study database. For this study, data on ROP were extracted from the medical charts, and when this information was lacking, local pediatricians were contacted for data on screening of ROP in the medical charts of their hospitals.

Cranial ultrasound was performed, as described previously, on 6 predefined occasions during the first week of life and were repeated 1 week and 1 month after birth. The most severe ultrasound abnormality, as assessed 4 weeks after birth, was used for classification. For the classification of ICHs, the classification of Volpe was used; for the classification of ischemic echodensities, a modified classification system according to Pidcock et al was used. We made a composite outcome of intracranial ultrasound findings as described by Scherjon et al.

Briefly, the criteria were as follows: normal, no ICH or a subependymal hemorrhage; echodensities less bright than the choroid plexus; suspect, intraventricular hemorrhage (<50% of lumen filled), any echodensities brighter than the choroid plexus, lasting <3 days; and abnormal, intraventricular hemorrhage (>50% of lumen filled) and any intraparenchymal hemorrhage, any echodensities brighter than the choroids plexus and lasting for >3 days. For each child, a socioeconomic score was available.

Assessments at 11 Years of Age on Visual Functioning

Visual Acuity, Visual Field, and Orthoptic Assessment
Ophthalmologic and orthoptic examinations were performed by experienced pediatric orthoptists (Drs Merckel and Everhard). The examination included the following:

1. Ophthalmologic history regarding visual acuity, refractive errors, and strabismus was obtained.
2. The corrected monocular and binocular visual acuity was determined. Acuity was expressed in Snellen acuity values assessed with Landolt-C-optotypes or with the Amsterdam Picture Chart. A visual acuity of ≥0.8 is considered normal.
3. Eye position and binocular function were assessed with the cover test and the alternating cover test at 30 cm and 2.5 m to categorize manifest or latent squint and esodeviation or exodeviation.
4. Stereoscopic vision was tested using the TNO random-dot stereo test.
5. Visual fields were tested with the Humphrey 91 screening test. When the children were not able to cooperate in Humphrey perimetry, the visual fields were assessed according to simple confrontation techniques as described by Donders. The final results were categorized in present or absent defects for each eye.

VMI Assessments
The visual motor assessment was performed by 2 experienced occupational therapists (Drs Verkerk and I. Hemmen). The assessment included the following tasks:

1. The Beery-Buktenica developmental test of VMI, in which increasingly complex geometric figures have to be copied with pencil on paper. The overall performance is converted into a standard score, based on chronological age, with a mean of 100 and an SD of 15. Results of >1 SD were considered as performing below or above age level.
2. The Motor Accuracy Test (MAT) was used to test the ability to trace directly over a printed black line using a red fine liner pencil with the preferred hand within 1 minute. The distance and the total length of the drawn line that is off the printed black line are measured and converted into scores, the accuracy, and the adjusted score with their additional SDs. The scores differ in that the adjusted score also includes the speed of the tracing. SD >0.9 is considered as performing below or above the age level. The results were extrapolated according to age, because norm data are available only until 10 years (mean age was 11.6 years).
3. The Motor-Free Visual Perception Test, Revised (MVPT-R) measures visual perceptual abilities, spatial relationships, visual discrimination, visual closure, visual memory, and figure-ground perception. A figure is presented and a matching item of 4 alternatives should be selected by pointing. The performance of the child is converted into a standard score that is based on the chronological age with a mean of 100 and an SD of 15. SD > 1 is considered as performing below or above the age level. For the first 3 tests, the child is sitting at a table; the tasks are presented at a distance of ~30 cm.
4. To get an impression of the ability to handle moving objects while standing, a subtest from the Movement Assessment Battery for Children (Movement ABC) was included. The ball skills offered were 2 tasks: to catch a ball and to throw a ball at a goal; the child had 10 opportunities for each task. The number of catches and hits was scored and compared with the test table of the Movement ABC, which is expressed in percentiles. A result of <5th percentile is considered to be below normal.

Statistical Analyses
Statistical analysis was performed with SPSS 11.5 for windows (SPSS, Chicago, IL). The primary outcome variables visual acuity, strabismus, visual field defects,
VMI, visual perception test (MVPT-R), MAT, and ball tasks were first analyzed in the total cohort (n = 98). In addition, outcome was analyzed in the 2 subgroups defined according to a normal (n = 58) or a raised (n = 31) U/C ratio. As test of significance, the χ² or linear by linear association was used as appropriate for dichotomous variables, whereas the t test or analysis of variance was used for continuous measurement.

All data were first analyzed using univariate analysis. After having detected the most significant variables (a significance level of P < .15 was used), logistic regression analysis was used to determine which variables influence the final outcome. Logistic regression was used to estimate the independent association of obstetric and neonatal variables with outcome parameters at the age of 11 years. As outcome variables, both visual motor function outcome parameters (normal or abnormal MAT, VMI, or MVPT-R results), and results from the ophthalmologic examination (normal or abnormal vision; strabismus or visual fields) were analyzed. The following explaining variables were included in the full model: U/C ratio, birth weight, use of antenatal corticoids, gestational age, gender, Apgar score at 5 minutes, asphyxia, sepsis, respiratory distress syndrome, intracranial abnormalities, bronchopulmonary dysplasia at the age of 28 weeks, and socioeconomic status. The inclusion in the model was partly based on differences between outcome groups as found by univariate testing, but possibly important variables (as known from the literature) were studied as well.

RESULTS

Ninety-eight children of the original cohort of 116 neonates were assessed at ~11 years of age, representing 85% of the survivors of the original cohort. Seven children could not be traced, and 9 parents refused to participate. Two children could not attend the follow-up session because they were living abroad. The 18 children who did not participate in the follow-up study showed similar perinatal characteristics as the assessed group (data not shown).

Perinatal characteristics of the children are given in Table 1. In 9 children, no antenatal Doppler measurements were performed. Mean U/C ratio in the group with brain sparing (n = 31) amounted to 1.71 ± 1.19 vs 0.43 ± 0.15 in the group without brain sparing (n = 58; P < .001). Fetuses with an abnormal U/C ratio were more often growth restricted as defined classically by growth curve characteristics. In 1989, in our hospital, antenatal corticosteroids were given for pregnancies with normal fetal growth only. Therefore, infants with a normal U/C ratio more often received antenatal corticoids. We did not find a statistically significant difference in the use of surfactant between the 2 groups: 6.9% in the normal U/C ratio group compared with 3.3% in the abnormal U/C ratio group (P = .47). However, at the time of the study period, surfactant was used only as a rescue treatment, in case of need of artificial ventilation with a fraction of inspired oxygen ≥60%. The incidence of intracranial ultrasound abnormalities, both severe ICHs and persisting echodensities as a sign of cerebral ischemia, was not significantly different between the 2 groups. The incidence of an intraventricular hemorrhage (grade 2 and higher) was for the normal and abnormal U/C group 15.5% and 9.7% (P = .44), respectively. The incidence of intraparenchymal echodensities that persisted for >3 days was not significantly different between the normal and the abnormal U/C ratio group: 17.2% and 6.5%, respectively (P = .156).

Because of early transfer to level 2 units, data on ROP screening were scarce. Therefore, we contacted the local pediatricians who cared for these infants after discharge from the NICU to provide data on ROP screening for this study. For 18 patients, no data on ROP were recorded; for 32 patients, neonatal charts had already been destroyed; in the other 48 patients, ROP screening had been documented. From the available data (50%), no

| TABLE 1 Perinatal Characteristics of the Children as Assessed at 11 Years of Age |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Characteristic                  | All Children (N = 98) | Normal U/C Ratio (n = 58) | Abnormal U/C Ratio (n = 31) | P (t Test or χ²) |
| Mean gestational age, wk        | 29.8 ± 1.9       | 29.4 ± 1.9       | 30.5 ± 1.8       | .009            |
| Mean birth weight, g            | 3130 ± 326       | 1378 ± 318       | 1137 ± 307       | .001            |
| Birth weight <10th percentile, %| 19 (19.4)        | 1 (1.7)          | 18 (58.1)        | .000            |
| Boys, %                         | 52 (53.1)        | 30 (51.7)        | 17 (54.8)        | .779            |
| Multiples, %                    | 38 (38.8)        | 21 (36.2)        | 10 (32.3)        | .636            |
| Apgar score <7, %               | 9 (9.2)          | 8 (13.8)         | 0 (0)            | .047            |
| Antenatal corticoids, %         | 68 (69.4)        | 51 (87.9)        | 12 (38.7)        | .000            |
| RDS, %                          | 22 (22.4)        | 15 (25.9)        | 3 (9.7)          | .070            |
| BPD (at 28 d), %                | 13 (13.3)        | 9 (15.5)         | 2 (6.5)          | .216            |
| Septicemia, %                   | 16 (16.3)        | 7 (12.1)         | 8 (25.8)         | .099            |
| Abnormal intracranial ultrasound, % | 22 (22.4) | 15 (25.9)        | 2 (6.5)          | .045            |

Data are mean ± SD or n (%). RDS indicates respiratory distress syndrome; BPD, bronchopulmonary dysplasia.

* Statistical significance between normal and abnormal U/C ratio group.

b For criteria for abnormal intracranial ultrasound, see “Methods.”
infants received a diagnosis of ROP grade 2 or more. No infants had any treatment for ROP.

At the follow-up visit at 11 years of age, anthropometric data were obtained in all children (Table 2). Only a slight difference in head circumference was found between the abnormal and normal U/C ratio group in favor of the children with a normal U/C ratio. This difference was not statistically different.

All children underwent orthoptic and ophthalmologic examination for assessment of visual acuity, visual fields, eye position, and binocular vision. The results are depicted in Table 3.

Five children in the normal U/C ratio group had a major impairment of visual acuity in 1 eye (<0.3). One child had a visual acuity of <0.3 in both eyes and was classified as having severe visual impairment. None of the children in the abnormal U/C group had a visual impairment.

Visual fields were tested with the Humphrey method in 90 children. Four children showed abnormalities of the visual fields. In these children, visual field defects were located in the nasal as well as the lower quadrants. In 8 cases (4 of the normal U/C ratio group, 2 of the abnormal U/C ratio group, and 2 of the group without antenatal Doppler measurement), the children were not able to perform the Humphrey perimetry, mainly because of severe motor and cognitive disabilities. In these children, the Donders confrontative method was used. In 3, visual field defects that were not located in specific quadrants were found.

Significantly more children with an abnormal U/C ratio had occlusion therapy during infancy because of amblyopia with or without strabismus. Of all children who had occlusion therapy during infancy, 70% had problems with binocularity. Of the 12 children with strabismus, 6 had esotropia and 6 had exotropia.

Visual motor testing could be performed in all children (Table 4). Twenty-six percent of all assessed children showed a VMI test below age level, but no significant difference was found between the U/C ratio groups. The disk-subset score as assessed at 5 years of age showed a fairly good correlation with the visual motor test at 11 years of age (r = 0.32; P < .01).

Also, very low scores were found on the MAT. Nearly half of the cohort performed below −0.9 SD, and the normal U/C ratio as well as the abnormal U/C ratio group showed similar results. Only 7 children scored higher than normal. Approximately 15% of the children had a low score on the MVPT-R. No statistically significant differences could be demonstrated between the U/C ratio groups.

A substantial percentage of the children (20%–30%) had problems with motor skills such as catching and throwing a ball as compared with the reference values. No differences were found between the U/C ratio groups.

Concerning visual motor function, logistic regression analysis revealed no clear independent association between certain explaining variables around birth and the studied outcome variables. MAT was associated with gestational age (P < .01), whereas VMI was associated with socioeconomic status (P < .03). MVPT-R was associated with asphyxia (P < .01) and intracranial abnormalities (P < .01).

Strabismus was associated with intracranial abnormalities (P < .01) as well as with asphyxia (P < .04). Strabismus is the only variable at 11 years that had an independent negative association with a high U/C ratio (P < .03). The other 2 ophthalmologic outcome parameters (visual acuity and visual fields) had no independent association with any of the explaining variables.

Table 5 shows the relation between the cerebral ultrasound findings in the neonatal period and visual outcome. The number of infants with an abnormal U/C ratio is significantly less in the group of abnormal cerebral ultrasound findings. Children with abnormalities on the neonatal ultrasound show significantly more impaired visual outcome in any of the tested domains.

**DISCUSSION**

**Visual Function in the U/C Ratio Groups**

In this study, we evaluated visual function at 11 years of age in a cohort of preterm infants with and without fetal brain sparing. Although we demonstrated differences in visual function at 1 and 5 years of age in this cohort in previous studies, we could not demonstrate any significant difference in visual functioning between the 2 groups at the age of 11 using extensive visual function testing.

The only difference was a higher percentage of children with a history of occlusion therapy in the group with fetal brain sparing compared with children without fetal brain sparing. It could be argued that the finding of the maturational differences of the visual evoked potentials in the first year of life and the opposite finding in VMI tests results at 5 years in our study group was a
finding by chance and not a real pathologic phenomenon. However, a similar cohort of children with and without fetal brain sparing were followed in Sweden. These investigators found an abnormal retinal vascular morphology in young adult life in the group with growth restriction, which they attributed to changes in fetal programming of infants with IUGR as described by Barker et al. This finding also indicates an association between an abnormal development in visual function and IUGR.

**Visual Function in the Whole Cohort**

The findings in the whole group are intriguing as compared with a normal population. This cohort showed a relatively high incidence of problems on visual functioning. We found a high percentage of children with strabismus (12%), whereas 6% of the children had severe visual impairment in 1 or both eyes, although no infant received a diagnosis of severe ROP. Binocular vision was absent in 12% of the children. Cooke et al conducted a similar study at the age of 7 years in children who were born at <32 weeks’ gestational age compared with matched term control infants and found similar percentages of poor visual acuity, strabismus, and stereoscopic vision in the preterm children.

Holmstrom et al reported an incidence of visual acuity abnormalities of 45%. We found a much lower incidence (17%) of visual acuity of 0.8 in 1 or both eyes. The cohort of Holmstrom et al consisted of a large group of infants with an extremely low gestational age: 60% of the infants were born at 29 weeks’ gestation; as a consequence, a high incidence (40%) of ROP was diagnosed. Unfortunately, we were not able to gather sufficient information on ROP screening in the neonatal period. It seemed that, irrespective of national guidelines for ROP screening, no careful documentation had been recorded. Also, one third of the medical charts had been destroyed after a period of 10 years. Therefore, we were

### Table 3

<table>
<thead>
<tr>
<th>Parameter</th>
<th>All Children (N = 98), n (%)</th>
<th>Normal U/C Ratio (n = 58), n (%)</th>
<th>Abnormal U/C Ratio (n = 31), n (%)</th>
<th>P (χ²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visual acuity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OD (≤0.8)</td>
<td>12 (12.2)</td>
<td>7 (12.1)</td>
<td>3 (9.7)</td>
<td>.73</td>
</tr>
<tr>
<td>OS (≤0.8)</td>
<td>15 (15.3)</td>
<td>9 (15.5)</td>
<td>4 (12.9)</td>
<td>.74</td>
</tr>
<tr>
<td>ODS (≤0.8)</td>
<td>6 (6.1)</td>
<td>4 (6.9)</td>
<td>0 (0)</td>
<td>.29</td>
</tr>
<tr>
<td>Perimetry or confrontive method</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual fields</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OD (abnormal)</td>
<td>6 (6.7)</td>
<td>3 (5.3)</td>
<td>3 (9.7)</td>
<td>.41</td>
</tr>
<tr>
<td>OS (abnormal)</td>
<td>4 (4.5)</td>
<td>2 (3.4)</td>
<td>2 (6.5)</td>
<td>.61</td>
</tr>
<tr>
<td>Strabismus and/or amblyopia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>History of occlusion</td>
<td>17 (17.3)</td>
<td>6 (10.3)</td>
<td>9 (29.0)</td>
<td>.04</td>
</tr>
<tr>
<td>Strabismus</td>
<td>12 (12.2)</td>
<td>5 (8.6)</td>
<td>5 (16.1)</td>
<td>.31</td>
</tr>
<tr>
<td>Absent stereoscopic vision</td>
<td>12 (12.2)</td>
<td>5 (8.6)</td>
<td>5 (16.1)</td>
<td>.31</td>
</tr>
</tbody>
</table>

OD indicates right eye; OS, left eye; ODS, right and left eyes.

* Humphrey perimetry result or, in case of disability, the result of the Donders method was analyzed.

### Table 4

<table>
<thead>
<tr>
<th>Test</th>
<th>All Children (N = 98), n (%)</th>
<th>Normal U/C Ratio (n = 58), n (%)</th>
<th>Abnormal U/C Ratio (n = 31), n (%)</th>
<th>P (t Test; χ²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beery-Buktenica (VMI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average performance</td>
<td>66 (67.2)</td>
<td>40 (68.9)</td>
<td>21 (67.7)</td>
<td>882</td>
</tr>
<tr>
<td>High performance</td>
<td>6 (6.1)</td>
<td>3 (5.2)</td>
<td>1 (3.2)</td>
<td></td>
</tr>
<tr>
<td>Low performance</td>
<td>26 (26.5)</td>
<td>15 (25.9)</td>
<td>9 (29.0)</td>
<td></td>
</tr>
<tr>
<td>Motor accuracy test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than −0.9</td>
<td>46 (46.9)</td>
<td>28 (48.3)</td>
<td>15 (48.4)</td>
<td>210</td>
</tr>
<tr>
<td>More than or equal to −0.9 and less than or equal to +0.9</td>
<td>45 (45.9)</td>
<td>29 (50.0)</td>
<td>13 (41.9)</td>
<td></td>
</tr>
<tr>
<td>More than +0.9</td>
<td>7 (7.1)</td>
<td>1 (1.7)</td>
<td>3 (9.7)</td>
<td></td>
</tr>
<tr>
<td>MVPT-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal score</td>
<td>73 (74.5)</td>
<td>46 (79.9)</td>
<td>21 (67.7)</td>
<td>343</td>
</tr>
<tr>
<td>High score</td>
<td>11 (11.2)</td>
<td>4 (6.9)</td>
<td>5 (16.1)</td>
<td></td>
</tr>
<tr>
<td>Low score</td>
<td>14 (14.3)</td>
<td>8 (13.8)</td>
<td>5 (16.1)</td>
<td></td>
</tr>
<tr>
<td>Catching a ball</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;5th percentile</td>
<td>22 (22.4)</td>
<td>11 (18.9)</td>
<td>8 (25.8)</td>
<td>236</td>
</tr>
<tr>
<td>Throwing a ball</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;5th percentile</td>
<td>31 (31.6)</td>
<td>19 (32.7)</td>
<td>9 (29.0)</td>
<td>617</td>
</tr>
<tr>
<td>Right handed</td>
<td>82 (83.6)</td>
<td>49 (84.5)</td>
<td>24 (77.4)</td>
<td>408</td>
</tr>
</tbody>
</table>

* Subtests of the Movement ABC.

These investigators found an abnormal retinal vascular morphology in young adult life in the group with growth restriction, which they attributed to changes in fetal programming of infants with IUGR as described by Barker et al. This finding also indicates an association between an abnormal development in visual function and IUGR.
not able to evaluate relationships of ROP screening results with final visual acuity.

In this cohort, strabismus was found in 12%. This is an ~3 times higher incidence than in a normal population. The incidence of strabismus is found to be 5% in normal preschool children in the Netherlands, and the incidence of amblyopia is 3% and of refractive errors is 5%.41 In other cohorts of preterm infants13,39,40,42 the incidence of strabismus varied between 13% and 22%.

We found a much higher incidence of strabismus (36%) in children with severe abnormal cerebral ultrasound findings; this was also found by others.9,16 The higher incidence of strabismus in preterm-born neonates seems to be related to cerebral lesions, because the incidence of strabismus in neonates without cerebral ultrasound defects (6%) is comparable to that in the normal population (5%).

We tested the visual fields with Humphrey perimetry and found abnormalities in ~5% of the children, mainly in the nasal fields. Using a Damato campimeter, O’Connor et al42 detected in just 1 child of 293 abnormal visual fields. Larsson et al43 assessed the visual fields by other methods. In their study, only preterm infants who were cryotreated for ROP showed constriction of the peripheral fields, whereas all preterm infants showed reduced neural capacity of the central fields.

Because a validated test on writing capacities is not available for this age group, we preferred to include a test on motor accuracy in the test battery. Remarkably, nearly 50% of the children scored below age level. It is possible that these subnormal scores have been biased by our need to extrapolate our findings for age. In addition, 26% of the children scored below age level on the VMI, and >20% of the children showed problems with the ball tasks of the Movement ABC. However, the results of MVPT-R were in agreement with the normal population.

It is interesting that a strong association with orthoptic and ophthalmologic examinations could not be detected. Similar results were also found in the study of Cooke et al.39

### Head Circumference

Although the abnormal U/C ratio group had a smaller head circumference compared with the normal U/C ratio group, the difference was not statistically significant (P < .07). In contrast to the recent findings by Cooke et al,39 we did not find any relationship between visual outcome test and head circumference (data not shown). Even after subdivision into normal and abnormal U/C ratio, no relation could be shown.

In this respect, it is important to mention that Tolsa et al44 demonstrated a significant reduction of intracranial volume and cerebral gray matter volume with a 3-dimensional MRI technique in infants with IUGR. In that cohort, however, infants were of a similar gestational age, whereas in our cohort, the abnormal U/C group had a higher gestational age compared with the normal group. This could have diminished the chance to detect a possible effect in our cohort.

### Abnormal Neonatal Cerebral Ultrasound Findings and Visual Function

Abnormal visual functioning at 11 years was strongly associated with an abnormal cerebral ultrasound result in the neonatal period. The children who developed cerebral hemorrhages and/or periventricular leukomalacia in the neonatal period showed bad visual outcomes on various domains, visual acuity, visual fields, strabismus, and visual motor function tests. Even the outcome of the MVPT-R, which is developed to test visual perception while avoiding motor function disorders, is much worse in the children with cerebral damage.
CONCLUSIONS
On the basis of these follow-up data up to the age of 5, we suggest that the brain-sparing effect is a benign adaptive mechanism that occurs in the fetus with IUGR. This extended follow-up study has confirmed the adaptive mechanism that occurs in the fetus with IUGR. In children with antenatal brain sparing, we found an accelerated maturation of visual evoked potentials during the first year of life, whereas visual motor function was impaired at 5 years of age. However, we could not demonstrate any difference in visual functioning at 11 years of age in these children.

ACKNOWLEDGMENTS
The study was supported by a grant from the Dutch Brain Foundation. The study was supported by a grant from the Dutch Brain Foundation.

REFERENCES
37. Hellström A, Dahlgren J, Marsál K, Ley D. Abnormal retinal vascular morphology in young adults following intrauterine growth restriction. Pediatrics. 2004;113(2). Available at: www.pediatrics.org/cgi/content/full/113/2/e77
38. Barker DJ, Bull AR, Osmond C, Simmonds SJ. Fetal and placental size and risk of hypertension in adult life. BMJ. 1990; 301:259–262
41. Biesheuvel-Snellen FMM. Appearances Also Count [in Dutch]. Rotterdam, Netherlands: Erasmus University; 1990
Auditory Brainstem Response Abnormalities and Hearing Loss in Children With Craniosynostosis

Michael W. Church, PhD, Leslie Parent-Jenkins, MS, Arlene A. Rozzelle, MD, Frances E. Eldis, PhD, S. Nadya J. Kazzi, MD

Departments of Obstetrics and Gynecology and Pediatrics, Wayne State University School of Medicine, Detroit, Michigan; Communication Disorders Clinic and Department of Plastic Surgery, Children’s Hospital of Michigan, Detroit, Michigan

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

ABSTRACT

OBJECTIVES. Craniosynostosis is a devastating disorder characterized by premature closure of the cranial plates before or shortly after birth. This results in an abnormally shaped skull, face, and brain. Little is known about hearing disorders in such patients, and nothing has been published about their auditory brainstem responses. Our objective was to evaluate such patients for auditory brainstem response and hearing disorders with the long-term goal of improving patient evaluation and management.

PATIENTS AND METHODS. We evaluated the auditory brainstem responses, hearing, and brain images of children with fibroblast growth factor receptor 2 craniosynostosis (n = 11).

RESULTS. Prolongation of the auditory brainstem response I-to-III interpeak latency was a frequent characteristic of fibroblast growth factor receptor 2 craniosynostosis, occurring in 91% of our patients. Prolongation of the III-to-V interpeak latency was an occasional characteristic, occurring in 27% of our patients. Whenever the I-to-III interpeak latency was prolonged, wave II was always abnormal. Associated morbidities included sensorineural hearing loss (27%), recurrent otitis media (100%), and Arnold-Chiari malformation (27%). Cranial decompression improved the interpeak latencies of 2 children.

CONCLUSIONS. These previously undocumented auditory brainstem response abnormalities reflect abnormal neural transmission, which could cause peripheral and central auditory processing disorders. We speculate that the major pathogenic basis of the I-to-III interpeak latency and wave II abnormalities is compression of the auditory nerve as it passes through the internal auditory meatus and posterior fossa, which would explain the auditory nerve hearing loss, tinnitus, and vertigo that affect these children. Awareness of these abnormalities could lead to important advancements in the auditory and neurosurgical assessment and management of this overlooked patient group. We provide recommendations for the improved assessment and management of these patients. In particular, we recommend that auditory brainstem response diagnostics become standard clinical care for this patient group as the best way to detect auditory nerve compression.
CRANIOSYNOSTOSIS is a birth defect characterized by premature closure of the cranial plates before or shortly after birth, resulting in a devastating condition wherein the child has an abnormally shaped skull, face, brain, low-set and posteriorly rotated ears, abnormal pinna configuration, eustachian tube dysfunction, osseous fixation, constricted and distorted brain growth, neurodevelopmental impairments or mental retardation, hydrocephaly, cranial nerve disorders, increased intracranial pressure, hearing loss, tinnitus, vertigo, facial palsy, seizures, and blindness. Treatment consists of surgery to relieve pressure on the brain and to remodel the child’s face and skull.1,2 Craniosynostosis occurs in ~1 of 2000 live births with >2000 such children being born in the United States every year.3 Most forms of craniosynostosis are caused by mutations in the fibroblast growth factor receptor (FGFR) 1, 2, or 3 genes, which influence bone and connective tissue growth. Our report involves auditory brainstem response (ABR), hearing, and brain imaging assessments of patients with the FGFR2 forms of craniosynostosis, which include Apert, Pfeiffer, Crouzon, Jackson-Weiss, and Beare-Stevenson syndromes.1,2

Research on hearing disorders in craniosynostosis patients is rather limited. Studies on Apert syndrome patients have described incidences of conductive hearing loss (CHL) because of recurrent otitis media (ROM) and congenital stapes fixation.4,5 There are no reports of ABR findings or sensorineural hearing loss (SNHL) in Apert syndrome patients. The literature on hearing loss in Pfeiffer syndrome is also limited. Investigators have found mild CHL and middle ear abnormalities,6 problems with middle ear effusion,7,8 and some instances of SNHL.9,10 There are no studies reporting ABR findings in Pfeiffer syndrome. The literature on hearing loss in Crouzon syndrome is even sparser. A case study found SNHL in one 13-year-old Crouzon child.11 Another study found instances of CHL (21%), mixed hearing loss (11%), and SNHL (21%).12 There are no reports of ABR or ROM findings in Crouzon patients. We found no literature on ABR or hearing disorders in Jackson-Weiss or Beare-Stevenson syndrome patients.

In testing an infant with Apert syndrome with ABRs, we noticed that he had abnormally prolonged wave I-to-III interpeak latencies (IPLs). We wanted to see if our finding was a common feature of craniosynostosis and if this ABR abnormality was associated with certain brain abnormalities and hearing losses. This article presents findings from a case series of FGFR2 craniosynostosis patients and discusses the clinical significance and possible pathogenic bases of our findings.

PATIENTS AND METHODS

All of the subjects (n = 11) were patients of the Communication Disorders Clinic, Children’s Hospital of Michigan. An informed consent, approved by the Wayne State University Human Investigation Committee, was obtained from a parent or legal guardian. Standard ABR recording procedures were used and collected on Biologic sensory evoked potential equipment using THD-39 headphones (Telephonics Corp, Farmington, NY) or insert earphones.13 Stimulus clicks (0.1 millisecond in duration, 23.3 clicks per second) were 75 dB of peak equivalent sound pressure level unless otherwise stated. Brain electrical activity (100- to 3000-Hz bandpass filter) was recorded by an active electrode placed on the upper forehead, a reference electrode on the earlobe of the stimulated ear, and a ground electrode on the left frontal area of the forehead.14 The ABR I-to-III and III-to-V IPLs for each patient were compared with published norms for infants and young children15-17 and norms for subjects ≥12 years of age.18 ABR latencies from normal-hearing children in our clinic compare favorably with these published norms. For example, normal-hearing 1-year-old children in our clinic had I-to-III IPLs of 2.32 ± 0.14 milliseconds (n = 20) as compared with the published norm of 2.31 ± 0.15 milliseconds (n = 47).15-17 The respective III-to-V IPLs were 2.04 ± 0.08 and 2.01 ± 0.22 milliseconds. Each patient’s IPLs were expressed as standard deviates (z scores) to determine their percentile ranking. IPLs that were in the upper 5th percentile (P ≤ .05) were considered to be significantly prolonged. Subjects received on multiple occasions a standard battery of age-appropriate hearing tests, including pure tone or behavioral audiometry, tympanometry, otoscopy, otoacoustic emissions, and sometimes ABR audiometry (latency-intensity profiles). Normal hearing was defined as −10 to +15 dB of hearing level across all of the frequencies.19,20 Standard criteria for defining ROM were used.19,20 Brain imaging included MRI and computed tomography scans.

RESULTS

Apert Case 1

This term black male infant had head circumference and body weight below the 10th percentile. Dysmorphic features included craniosynostosis, polydactyly, and syndactyly of the feet. The ABRs, obtained at 8 days postpartum, had significant prolongations of the wave I-to-III IPLs. The left- and right-ear I-to-III IPLs of 3.90 and 3.84 milliseconds were prolonged by 4.50 and 4.27 SDs in comparison with the age-matched norm (mean ± SD) of 2.704 ± 0.266 milliseconds (P < .001). In contrast, this infant’s left- and right-ear wave III-to-V IPLs of 2.64 and 2.76 milliseconds were within normal limits of the age-matched norm of 2.379 ± 0.246 milliseconds.

Subsequent brain imaging indicated an Arnold-Chiari (A-C) malformation with herniated cerebellar tonsils and enlarged lateral ventricles but “with no evidence of brainstem morphologic abnormality.” Cranial surgery was done at 22 months of age to relieve intracranial...
pressure. During recovery, he was given another ABR series, which indicated that he continued to have abnormally prolonged I-to-III IPLs. Specifically, the I-to-III IPLs from the left and right ears were 2.83 and 3.07 milliseconds (Fig 1), which were significantly prolonged by 3.21 and 4.38 SDs in comparison with the age-matched norm of 2.168 ± 0.206 milliseconds \((P < .001)\). Wave II was absent in both ears. In contrast, this child’s left- and right-ear III-to-V IPLs of 1.89 and 1.71 milliseconds compared favorably with the age-matched norm of 1.959 ± 0.195 milliseconds. A comparison of the left ear presurgery and postsurgery I-to-III IPLs indicates the prolongation decreased from 4.50 to 3.21 SDs. Thus, there was an improvement in the left-ear I-to-III IPL after decompression surgery. In contrast, the right-ear I-to-III IPL showed no improvement. Hearing tests indicated no SNHL or permanent CHL.

**Apert Case 2**

Brain imaging of this black female infant showed enlarged lateral ventricles. She had a history of ROM with no SNHL or permanent CHL. ABRs were collected at 6 months of age. The 2.89- and 2.83-millisecond I-to-III IPLs for the right and left ears were prolonged by 2.11 \((P = .017)\) and 1.84 SDs \((P = .033)\), respectively. Visual inspection suggested that wave II latency was also prolonged in both ears. In contrast, the III-to-V IPLs for both ears were within normal limits. Fig 2 shows the right-ear ABR. The left-ear ABR was similar.

**Apert Case 3**

This white male has brachycephaly and plagiocephaly, enlarged third and fourth ventricles, and required a ventricular shunt. Brain imaging indicated no A-C malformation at 1 month of age, but by 29 months of age, a Chiari I malformation had developed. He is globally delayed and has central sleep apnea. He had a moderate mixed hearing loss (SNHL + CHL) in the right ear as evidenced by ABR audiometry, normal hearing in the left ear, and a history of ROM. ABRs were collected at 22 months. The 2.77- and 2.83-millisecond I-to-III IPLs for the left and right ears were prolonged by 2.92 and 3.21 SDs, respectively \((P < .001)\). In contrast, the III-to-V IPLs were within normal limits. Fig 3 shows the right-ear ABRs. The left-ear ABRs were highly similar. Wave II was absent in both the left- and right-ear ABR traces.

**Apert Case 4**

Brain imaging of this white male infant indicated shortening of the posterior fossa, elongation of the middle cranial fossa, a foreshortened anterior cranial fossa, moderate enlargement of the lateral ventricles, and no A-C malformation. The Bayley Scales of Infant Development found delayed motor (1st percentile) and cognitive skills (5th percentile). He had mild-to-moderate CHL because of scarring from repeated myringotomies and ROM. ABRs were collected at 3 years of age. The 2.95-millisecond I-to-III IPL for the left ear was prolonged by 3.95 SDs \((P < .001)\), whereas the III-to-V IPL was within normal limits. The left-ear ABR also showed poor morphology and no wave II. In contrast, both the I-to-III and III-to-V IPLs for the right ear were within normal limits, the ABR had excellent morphology, and wave II was normal and clearly defined (Fig 4).

**Apert Case 5**

This 26-year-old black man is the father of Apert case 1. He has a history of ROM but otherwise normal hearing. His right-ear ABR had I-to-III and III-to-V IPLs of 3.07 and 2.54 milliseconds, which were 6.40 and 4.86 SDs above the respective norms of 2.11 ± 0.15 and 1.86 ± 0.14 milliseconds \((P < .001)\). The right-ear wave II was not clearly defined (Fig 5). Left-ear ABRs were not possible in this patient because of impacted earwax. Brain images were not available.

**Apert Case 6**

This 16-year-old white male has mild bilateral high frequency SNHL, a history of ROM, attention-deficit/hyperactivity disorder, language delay, a slightly deformed
posterior fossa, and an arachnoid cystic area posterior to the cerebellar vermis. This teenager had right-ear ABR I-to-III and III-to-V IPLs that were within normal limits, and the wave II was clearly defined (Fig 6). This is our only patient with all of the IPLs within normal limits.

Apert Case 7
This black male infant, born after 36 weeks of gestation, had meconium aspiration and received mechanical ventilation for 6 days. Brain imaging indicated decreased anterior-posterior diameter, midface hypoplasia, mild enlargement of the lateral ventricles, a cavum septum pellucidum and vergae, prominent Virchow-Robin spaces bilaterally, and no A-C malformation. He had normal hearing in both ears. He had an ABR test at 2.5 months of age. The left-ear I-to-III IPL of 3.42 milliseconds was prolonged by 4.17 SDs (P = .001). Wave II was poorly defined in both the left- and right-ear ABRs. In contrast, the left-ear III-to-V IPL of 2.48 milliseconds was just barely within the upper range of normalcy, being prolonged by 1.64 SDs (P = .0505; Fig 7). A temporary CHL in the right ear, because of middle ear fluid, prevented the right-ear IPLs from being scored.

Pfeiffer Case 1
Brain imaging of this white male infant indicated a small brainstem, malformed posterior fossa, a normal fourth ventricle with malformed third and lateral ventricles, suspected aqueduct stenosis, absent septum pellucidum, cortical dysplasia, optic nerve and chiasm hypoplasia, pituitary hypoplasia, and no A-C malformation. He had normal hearing in both ears. He was ABR-tested at 3 months of age during recovery from anesthesia for his presurgery brain imaging. The I-to-III IPLs of 3.48 and
3.30 milliseconds for the left and right ears were prolonged by 4.45 and 3.61 SDs, respectively ($P < .001$). The III-to-V IPL of 2.36 milliseconds for the left ear was within normal limits. In contrast, the right-ear III-to-V IPL of 2.60 milliseconds was prolonged by 2.20 SDs ($P = .014$). Wave II was essentially absent from both the left- and right-ear ABR traces (Fig 8).

Crouzon Case 1
This white male infant, aged 12 months, had mild lateral ventricle asymmetry and ventriculomegaly, a very distorted cortex, myelination appropriate with chronologic age, and distortion of the posterior fossa with the brainstem shifted to the left of midline, suggesting that the right-ear auditory nerve was stretched. The cranio-cervical junction showed a C1 ring closely juxtaposed to the occiput and a rotary deformity at the C1 to C2 level. The skull base angle was quite distorted and measured $13^\circ$. This child had normal hearing. The left-ear ABR had I-to-III and III-to-V IPLs that were within normal limits. In contrast, the right-ear I-to-III IPL of 3.60 milliseconds was prolonged by 8.64 SDs ($P < .001$), and the wave II was distorted. The right-ear III-to-V IPL was within normal limits (Fig 9).

Crouzon Case 2
This white female subject, aged 15 years, had mild lateral ventricular dilation with unusual configuration of the frontal horns but no dilation of the other ventricles, slightly smaller right cerebellar hemisphere, asymmetry of both cortices, brain parenchyma showing good gray-white matter differentiation with no focal areas of abnormality, no gross intracranial abnormalities, midface and nasal septum deviated posteriorly to the left, and mild calvarial asymmetry with left frontal and right occipital plagiocephaly. She had mild-moderate SNHL $\leq 2$ kHz in the left ear and mild SNHL $\leq 2$ kHz in the right ear. She had no permanent CHL, a history of ROM, and persistence of otitis media into adolescence. This patient’s wave IV/V complex was dominated by the wave...
IV peak with the wave V appearing as a shoulder on the downslope of the wave complex. The left-ear I-to-III IPL of 2.30 milliseconds was within the normal limits of 2.11 ± 0.15 milliseconds, whereas the III-to-V IPL of 2.30 milliseconds was prolonged by 3.14 SDs (P = .002) compared with the age-matched norm of 1.86 ± 0.14 milliseconds. The reverse situation occurred in the right ear, where the I-to-III IPL of 2.54 milliseconds was prolonged by 2.87 SDs (P = .002), and the III-to-V IPL of 2.06 was within normal limits. The left- and right-ear I-to-V IPLs of 4.60 milliseconds were prolonged by 3.00 SDs (P = .001) compared with the age-matched norm of 3.94 ± 0.22 milliseconds. Wave II was prolonged in the right ear with a latency of 3.21 milliseconds compared with a left-ear latency of 2.80 milliseconds, despite normal wave I latencies of 1.62 and 1.68 milliseconds, respectively (Fig 10).

**Jackson-Weiss Case 1**
This black male child, aged 4 years 2 months, had mild sleep apnea, midface retrusion, hypertelorism, exophthalmia, left lateral ventriculomegaly, a thin and dysmorphic corpus callosum, a small posterior fossa, mildly diminished cerebrospinal fluid flow at the posterior aspect of the foramen magnum, an A-C malformation, and markedly abnormal calvarial shape. The cervical spine had expansion of the spinal cord from C2 to C7 with septations. Mild subluxation was seen between C1 and C2. There was incomplete fusion of the anterior arch of the C1 ring. He had a visual loss of unknown origin in the left eye. He had a history of ROM but otherwise normal hearing.

He had a Chiari decompression surgery before his first ABR test. For this first ABR test, the left-ear I-to-III IPL of 3.01 milliseconds was prolonged by 5.93 SDs (P < .001), and the right-ear I-to-III IPL of 2.42 milliseconds was prolonged by 2.00 SDs (P = .023). In contrast, the left-ear III-to-V IPL of 1.59 milliseconds was actually significantly shorter than normal by −2.29 SDs (P = .011), and the right-ear III-to-V IPL of 1.95 milliseconds was within normal limits. Wave II was absent in both ears (Fig 11). This child had a repeat Chiari decompression surgery followed by another ABR test. The left- and right-ear I-to-III IPLs of 3.07 and 2.48 milliseconds were essentially unchanged, whereas the left- and right-ear III-to-V IPLs of 1.06 and 1.65 milliseconds were now noticeably shorter by 0.47 and 0.30 milliseconds, respectively.

**Summary**
Table 1 illustrates that prolonged I-to-III IPLs, SHNL, and ROM are common characteristics of FGFR2 craniosynostosis. There was a frequent occurrence of left-right asymmetry with the I-to-III and III-to-V IPL prolongations.
and the SNHL, which were often worse in 1 ear. The occurrence and severity of I-to-III IPL prolongations did not always correspond with the presence and severity of SNHL or the presence of an A-C malformation.

**DISCUSSION**

Prolongation of the ABR I-to-III IPL was a frequent characteristic of our FGFR2 craniosynostosis patients, occurring in 10 (91%) of 11 patients. Prolongation of the III-to-V IPL was an occasional characteristic, occurring in 3 (27%) of 11 patients. Prolongation of the I-to-III IPL signifies slowed neural transmission times between the distal portion of the auditory nerve (wave I) and the cochlear nucleus (wave III). Prolongation of the III-to-V IPL indicates slowed neural transmission times between the cochlear nucleus and the lateral lemniscus/inferior colliculus (wave V).21 Wave II (proximal portion of the auditory nerve) was absent, prolonged, or dysmorphic in all 16 of the ears (100%) with prolonged I-to-III IPLs, suggesting that this portion of the auditory nerve was frequently compromised. Thus, these patients can have both upper and lower brainstem and auditory nerve dysfunctions. Our patients frequently had asymmetric ABR results. This is consistent with the asymmetric craniofacial features, brain morphology, and cranial nerve disorders that occur in these patients. The prolonged IPLs improved in 2 patients after posterior fossa decompression.

**Neurologic/Neurosurgical Implications**

Although prolonged IPLs have never been described in craniosynostosis patients before, IPL prolongations have been observed in patients with either A-C malformations or vascular compression of the auditory nerve. These 2 groups provide clues to the pathogenic origins of the IPL prolongations in our craniosynostosis children.

**A-C Malformation and the ABR**

I-to-III IPL prolongations are sometimes seen in A-C malformation patients. Increased posterior fossa pressure and brainstem deformation from cerebellar herniation are likely causes of their I-to-III IPL prolongations.22-24 Three (27%) of our 11 craniosynostosis patients had A-C malformations. Our patients fell into 3 groups: (1) 1 patient (9%) had normal I-to-III IPLs and no A-C malformation; (2) 3 patients (27%) had both prolonged I-to-III IPLs and A-C malformation; and (3) 7 patients (64%) had prolonged I-to-III IPLs but no A-C malformation. Thus, most of our patients had prolonged I-to-III IPLs in the absence of an A-C malformation. One

---

**TABLE 1**

Summary of ABR Abnormalities and Hearing Losses in FGFR2 Craniosynostosis

<table>
<thead>
<tr>
<th>Patient</th>
<th>I–III IPL Prolongation</th>
<th>III–V IPL Prolongation</th>
<th>SNHL</th>
<th>CHL</th>
<th>ROM</th>
<th>A–C Malformation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LE</td>
<td>RE</td>
<td>LE</td>
<td>RE</td>
<td>LE</td>
<td>RE</td>
</tr>
<tr>
<td>Apert 1</td>
<td>3.21</td>
<td>4.38</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Apert 2</td>
<td>2.11</td>
<td>1.84</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Apert 3</td>
<td>2.92</td>
<td>3.21</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>+</td>
</tr>
<tr>
<td>Apert 4</td>
<td>3.95</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>+</td>
</tr>
<tr>
<td>Apert 5</td>
<td>CND</td>
<td>6.40</td>
<td>CND</td>
<td>4.86</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Apert 6</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Apert 7</td>
<td>4.17</td>
<td>CND</td>
<td>—</td>
<td>CND</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Pfeiffer 1</td>
<td>4.45</td>
<td>3.61</td>
<td>—</td>
<td>2.20</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Crouzon 1</td>
<td>—</td>
<td>8.63</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Crouzon 2</td>
<td>—</td>
<td>3.14</td>
<td>2.87</td>
<td>—</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>Jackson-Weiss 1</td>
<td>5.93</td>
<td>2.00</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

IPL indicates interpeak latency; CND, could not determine; LE, left ear; RE, right ear; +, mild; ++, moderate; TYD, too young to determine; —, within normal limits.

*Values indicate the number of SDs above the age-matched norm (P ≤ .05).
patient had normal brainstem morphology despite an A-C malformation. One infant had a stretched auditory nerve that likely caused his I-to-III IPL prolongation by making neurotransmission time longer. One patient showed a moderate improvement (shortening) in his left-ear I-to-III IPL, and a second patient showed shortening of the III-to-V IPLs in both ears after posterior fossa decompression surgery. However, their postsurgical I-to-III IPLs were still abnormally prolonged, and their wave IIs were still absent. So whereas A-C malformation, increased intracranial pressure, deformed brainstem, and stretched auditory nerve can be involved in I-to-III IPL prolongation, they did not explain most of our cases.

Vascular Compression and the ABR
Vascular compression of cranial nerves in the cerebellopontine angle (CPA) can cause facial spasms, auditory nerve hearing loss, vertigo, tinnitus, and other disorders. Patients with vascular compression along the auditory nerve have I-to-III IPL prolongations and absent or distorted wave IIs. Microvascular decompression (MVD) surgery is often successful in resolving or ameliorating the hearing loss, vertigo, tinnitus, facial spasms, and other CPA symptoms, as well as improving the patients’ ABR abnormalities. Thus, auditory nerve compression merits consideration as a major source of the I-to-III IPL prolongations and absent and/or dysmorphic wave IIs in our craniosynostosis patients.

Pathogenic Bases
Although we analyzed a small population, we can make several generalizations about the pathogenic bases and neurologic implications of our ABR findings. First, A-C and brainstem malformations or stretched auditory nerves contribute to the abnormally prolonged I-to-III IPLs in some craniosynostosis patients. However, we speculate that the major pathogenic sources of the I-to-III IPL prolongations are the vascular, bony, and connective tissue problems that often afflict such patients around the skull, cranial nerves, and brainstem. In particular, the high prevalence of wave II abnormalities in our patients strongly suggests a compression of the auditory nerve by vascular, bony, or connective tissues as it passes through the internal auditory meatus and/or the posterior fossa. Second, the auditory nerve, facial nerve, and internal auditory branch of the basilar artery transmit collectively through the internal auditory meatus and posterior fossa. The causes of deafness, tinnitus, vertigo, and hemifacial spasms in craniosynostosis patients are currently unknown. Thus, one implication from our study is that auditory and facial nerve compressions are the major pathogenic sources of these morbidities. If true, I-to-III IPL prolongations and wave II abnormalities could be important predictors of auditory nerve hearing loss, tinnitus, vertigo, hemifacial spasms, and reduced basilar artery perfusion. Third, MVD surgery could preserve or restore the integrity of the auditory and facial nerves, the basilar artery, and associated brainstem functions.

Neurologic and Neurosurgical Recommendations
We recommend that craniosynostosis patients undergo ABR evaluation with separate testing for each ear, that they be periodically evaluated with ABRs during their life span to provide early detection of developing neurologic dysfunction that may not be apparent by brain imaging or other evaluations, that patients with wave II abnormalities be assessed for facial nerve and basilar artery disorders, that the region around the internal auditory meatus be examined for possible surgical intervention, and that ABRs be gathered both presurgically and postsurgically to assess the need, extent, and success of cranial, posterior fossa, and MVD surgeries. It is likely that the severity of an IPL prolongation reflects the degree of brainstem and cranial nerve pathology and should, therefore, be useful in predicting clinical signs of cranial nerve, brainstem, and CPA syndromes. Assessing craniosynostosis patients for I-to-III and III-to-V IPL prolongations could thereby aid in the decision to have neurosurgery, particularly when neurologic signs are subtle or progressing.

Hearing Implications
Three (27%) of 11 craniosynostosis patients had SNHL. This is comparable to the findings of others. Of the 20 ears with ABR results, our data fell into 4 groups regarding SNHL. First, 2 (10%) of 20 ears had neither SNHL nor prolonged I-to-III IPLs and were, therefore, ostensibly normal. Second, 3 (15%) of 20 ears had SNHL in the absence of I-to-III IPL prolongation. The source of SNHL in these ears was purely sensory (cochlear) in origin. Third, 13 (65%) of 20 ears had no SNHL but still had prolonged I-to-III IPLs. These ears likely had a purely neural abnormality located at the proximal portion of the auditory nerve. Fourth, 2 (10%) of 20 ears had both SNHL and prolonged IPLs. The pathogenic basis of the SNHL in these ears could be auditory nerve or cochlear dysfunction or both. Some patients with A-C malformation, after posterior fossa decompression surgery, experienced recovery from SNHL, as well as improvements in the ABR I-to-III IPL. Thus, increased intracranial pressure can be a pathogenic factor in some cases of SNHL.

Two patients (18%) had permanent CHL. The CHL in craniosynostosis patients can be caused by internal auditory canal abnormalities, stapedial fixation, Eustachian tube dysfunction, tympanic membrane pathology, or osicular erosion secondary to ROM and repeated myringotomies. In our patients, both cases of permanent CHL were because of tympanic membrane scarring secondary to ROM and repeated myringotomies. None had
internal auditory canal abnormalities or stapedia fixation. Previous studies of craniosynostosis patients report a high prevalence of CHL, but this seems to be primarily from temporary otitis media episodes rather than permanent middle ear pathology. Of the 7 patients old enough to have adequate otitis media histories, all (100%) had ROM, and all of the older patients showed a persistence of ROM into late childhood and adolescence. Two pathogenic sources of ROM in craniofacial syndrome children are immune deficiencies and Eustachian tube dysfunction.19,20

We can make some generalizations about the hearing implications of our findings:

1. The I-to-III and III-to-V IPL prolongations suggest auditory nerve and lower and upper brainstem dysfunctions. These dysfunctions can cause auditory nerve and central auditory processing disorders (CAPDs). CAPD can cause or exacerbate learning disorders, attention-deficit/hyperactivity disorder, and language delays and can compromise sound localization, auditory timing, and speech comprehension in the presence of competing sounds. An IPL prolongation may, therefore, predict the presence and severity of CAPD and its comorbidities. We recommend that craniosynostosis patients be assessed and managed accordingly.

2. SNHL and I-to-III IPL prolongation do not always coexist in craniosynostosis patients. The SNHL in these patients can be either cochlear or auditory nerve in origin. Testing should be done to determine whether the SNHL is cochlear or neural in origin and to manage the patient accordingly (eg, hearing aid or MVD).

3. Craniosynostosis patients will very likely have asymmetric SNHL and CHL in addition to asymmetric IPLs.

4. For those with permanent hearing losses, hearing aids were recommended for only 1 patient. Others did not need them, because their hearing losses were unilateral or <25 dB in the amplifiable range.

5. A-C malformation patients can show progressive hearing loss with aging and a persistence of pediatric ROM into adolescence.5-8,23 A high incidence and a persistence of ROM into adolescence are common in craniofacial syndromes.19,20,32,33 Therefore, craniosynostosis patients should be assessed throughout their life span to provide early detection and management of progressive hearing loss and the persistence of ROM beyond childhood.

Unexplained Observations

We had 2 puzzling observations for which we have no explanation: (1) one patient had a III-to-V IPL that was significantly shorter than normal, and (2) despite prolonged I-to-III IPLs and absent/dysmorphic wave IIs, the amplitudes and morphologies of waves III and V were relatively unaffected in many instances. The same is true for the vascular compression patients studied by M. B. Moller, MD, PhD, and A. R. Moller, PhD (written communication, 2006).

CONCLUSIONS

The ABR abnormalities of prolonged IPLs and absent/dysmorphic wave IIs in our craniosynostosis patients are previously undocumented observations that could change our perception of the disease process and lead to important clinical and research advances. For example, we hypothesize that the major pathogenic basis of these ABR abnormalities is a compression of the auditory nerve as it passes through the internal auditory meatus and posterior fossa. We further speculate that auditory and facial nerve compression and reduced basilar artery blood flow underlie the morbidities of auditory nerve deafness, tinnitus, vertigo, hemifacial spasms, and brainstem disorders. MRI tractography or functional brain imaging may be needed to investigate this nerve compression hypothesis, because conventional MRI was not able to image these nerves. We speculate that MVD surgery could prevent or reverse these cranial nerve morbidities and greatly improve the quality of life for these patients. We hypothesize that brainstem compression, A-C malformation, and a stretched auditory nerve can be occasional pathogenic sources of prolonged I-to-III and III-to-V IPLs and that posterior fossa decompression can ameliorate these conditions. The ABR should, therefore, be useful in evaluating the need and appraising the benefits of cranial, posterior fossa, and MVD surgery. We further hypothesize that craniosynostosis patients will have an inordinate prevalence of auditory nerve and CAPDs. The presence of these disorders should be investigated and managed. Finally, the ABR may be the best way to diagnosis auditory and facial nerve compression in craniosynostosis children. Others have concluded that the ABR is more useful than brain imaging for detecting vascular compression of the auditory nerve in tinnitus, vertigo, and hemifacial spasm patients (A. R. Moller, PhD, written communication, 2006). Thus, ABR diagnostics should be a standard part of neurologic, neurosurgical, and hearing evaluations for all craniosynostosis and other patients at risk for auditory nerve, facial nerve, and brainstem compression.

ACKNOWLEDGMENTS

We thank Karen Piggott, AUD, and Smita Somne, MS, for collecting data on 2 patients, as well as Seetha Shankaran, MD (Director, Neonatal/Perinatal Medicine, Children’s Hospital of Michigan) for reviewing our article.

REFERENCES

11. Vallino-Napoli LD. Audiologic and otologic characteristics of Crouzon syndrome.
One-Year Follow-up of Very Preterm Infants Who Received Lucinactant for Prevention of Respiratory Distress Syndrome: Results From 2 Multicenter Randomized, Controlled Trials

Fernando Moya, MD, Sunil Sinha, MD, PhD, Janusz Gadzinowski, MD, PhD, Ralph D’Agostino, PhD, Robert Segal, MD, Carlos Guardia, MD, Jan Mazela, MD, PhD, Genzhou Liu, PhD, on behalf of the SELECT and STAR Study Investigators

ABSTRACT

BACKGROUND. The benefits of exogenous surfactants for prevention or treatment of respiratory distress syndrome are well established, but there is a paucity of long-term follow-up data from surfactant-comparison trials.

OBJECTIVE. We sought to determine and compare survival and pulmonary and neurodevelopmental outcomes through 1 year corrected age of preterm infants who received lucinactant and other surfactants in the SELECT (Safety and Effectiveness of Lucinactant Versus Exosurf in a Clinical Trial) and STAR (Surfaxin Therapy Against Respiratory Distress Syndrome) trials individually and, secondarily, from analysis using combined data from these 2 trials.

METHODS. All infants from both trials who were randomly assigned to administration of lucinactant (175 mg/kg), colfosceril palmitate (67.5 mg/kg), beractant (100 mg/kg), or poractant alfa (175 mg/kg) were prospectively followed through 1 year corrected age, at which point masked assessment of outcomes was performed for surviving infants. One-year survival was a key outcome of interest. Other parameters assessed included rates of rehospitalization and respiratory morbidity and gross neurologic status. Data were analyzed by comparing the different surfactants within each trial and, in secondary analysis, combining data from both trials to compare lucinactant versus the animal-derived surfactants (beractant and poractant) used in these trials. Survival rates over time were compared by using the Wilcoxon test for survival through 1 year corrected age and logistic regression for comparison of fixed time points. The latter analyses were performed by using the prespecified approach, where loss to follow-up or withdrawal of consent was not considered as an event.
imputed as a death, and also using raw data. Other outcomes were analyzed by using the Cochran-Mantel-Haenszel test or logistic regression for categorical data, and analysis of variance on ranks was used for continuous data.

RESULTS. Very few cases were lost to follow-up in either trial (29 of 1546 enrolled in both trials [1.9%]). In the primary analysis of the SELECT trial comparing lucinactant to either colfosceril or beractant, there were no significant differences in the proportion of infants who were alive through 1 year corrected age. Fixed-time-point estimates of mortality at 1 year corrected age imputing loss to follow-up as a death were 28.1% for lucinactant, 31.0% for colfosceril, and 31.0% for beractant. By using raw data without imputing loss to follow-up as a death, mortality estimates at 1 year corrected age were computed to be 26.6%, 29.1%, and 28.3%, respectively. In the primary analysis of the STAR trial, significantly more infants treated with lucinactant were alive through 1 year corrected age compared with those who received poractant alfa. Fixed time estimates of mortality at 1 year corrected age imputing loss to follow-up as a death were 19.4% for lucinactant and 24.2% for poractant. These estimates using raw data that did not impute loss to follow-up as a death were 18.6% and 21.9%, respectively. In the combined analysis, survival through 1 year corrected age was higher for infants in the lucinactant group versus that of the infants in the animal-derived surfactants (beractant and poractant) group. The fixed-time-point estimates of mortality at 1 year corrected age imputing loss to follow-up as a death for lucinactant and animal-derived surfactants were 26.0% and 29.4%, respectively. However, the 1-year-corrected-age estimates using combined raw data were 24.6% for the lucinactant group and 26.7% for the animal-derived surfactant group. The incidence of post-discharge rehospitalizations, total number of rehospitalizations, incidence of respiratory illnesses, and total number of respiratory illnesses were generally similar among those in the treatment groups. Neurologic status at 1 year corrected age was essentially similar between infants who received lucinactant and those who received all other surfactants used in these 2 trials.

CONCLUSIONS. Findings from this 1-year follow-up of both lucinactant trials indicate that this new peptide-based synthetic surfactant is at least as good, if not superior, to animal-derived surfactants for prevention of respiratory distress syndrome and may be a viable alternative to animal-derived products.

INTRATRACHEAL ADMINISTRATION OF animal-derived and synthetic exogenous surfactant preparations improves respiratory status and decreases mortality and morbidity rates among premature infants at risk of or with respiratory distress syndrome (RDS). Currently available animal-derived surfactants from bovine or porcine sources contain phospholipids and variable, yet relatively small, quantities of surfactant proteins (SPs) B and C, whereas currently available synthetic surfactants contain phospholipids but no SPs. Although these synthetic surfactants have potential safety advantages over animal-derived products, they seem to be inferior to animal-derived surfactants in improving clinical outcomes.

A meta-analysis of 11 controlled trials that compared these 2 classes of surfactants demonstrated a marginally significant lower mortality rate and a lower risk of pneumothorax with animal-derived surfactants when surfactant was administered as a rescue therapy. However, no reductions in the incidence of bronchopulmonary dysplasia (BPD) were demonstrated. None of these surfactant-comparison trials reported findings beyond the initial hospital stay in the NICU. The limitation of synthetic non–protein-containing surfactants has been attributed to the lack of SP-B and SP-C. The absence of SP-B seems to be particularly important: animals or humans lacking SP-B because of a genetic mutation develop a fatal form of respiratory failure during the neonatal period. In contrast, individuals with mutations in SP-C develop interstitial lung disease as adults rather than neonatal RDS.

Lucinactant (Surfaxin; Discovery Laboratories, Inc, Warrington, PA) is a new-generation synthetic surfactant that contains phospholipids and a high concentration of the synthetic 21-amino acid hydrophobic peptide (sinapultide, also known as KL4 peptide). This peptide resembles the hydrophobic-hydrophilic amino acid pattern of the tail end of SP-B. The concentration of sinapultide in lucinactant is higher than the concentration of SP-B in current animal-derived products, which approximates the concentration of SP-B in normal lungs. Lucinactant has greater resistance to oxidation and per-oxidation than the bovine-derived surfactant beractant (Survanta; Ross Products Division, Abbott Laboratories, Columbus, OH) and has been shown to improve pulmonary function and alveolar expansion in an animal model of RDS and in a pilot study that involved preterm infants with RDS.

We recently reported the results of 2 multicenter, phase III, double-blind, randomized, controlled trials, which demonstrated the efficacy of lucinactant in the prevention of neonatal RDS. The SELECT (Safety and Effectiveness of Lucinactant Versus Exosurf in a Clinical Trial) study compared lucinactant with the synthetic non–protein-containing surfactant, colfosceril palmitate (Exosurf; GlaxoSmithKline, Brentford, United Kingdom); beractant was used in the trial as a reference arm. The STAR (Surfaxin Therapy Against Respiratory Distress Syndrome) trial compared lucinactant with the porcine-derived surfactant poractant alfa (Curosurf;
Chiesi Farmaceutici, Parma, Italy).12 Inclusion criteria, approach to surfactant administration, and time periods when the studies were conducted were fairly similar for both trials. In the SELECT trial, RDS-related mortality through 14 days of age was significantly (\( P < 0.01 \)) reduced with lucinactant compared with both beractant and colfosceril palmitate, and proportionally more patients were alive at 36 weeks’ postmenstrual age (PMA) compared with beractant (odds ratio [OR]: 0.67; 95% confidence interval [CI]: 0.45–1.00). In the STAR trial, the primary outcome of alive without BPD at 28 days was not significantly different between lucinactant and poractant alfa.12 To evaluate outcomes beyond the initial hospital stay and examine further the safety of lucinactant, both trials included planned follow-up of participating infants to 1 year corrected age.

The primary objective for this study was to report the outcome results of the planned follow-up to 1 year corrected age of infants participating in the SELECT and STAR trials. Furthermore, given the similarity of these trials in the populations studied, treatment approach, end points, and contemporary nature, a secondary goal of this analysis was to compare the outcome of infants who received lucinactant versus those who received other classes of surfactants after combining data from both trials.

**METHODS**

**Study Design**

Methods for both RDS-prevention trials have been described previously in detail.11,12 In summary, within 30 minutes of birth, infants between 600 and 1250 g were randomly assigned to receive lucinactant, colfosceril palmitate, or beractant in a 2:2:1 ratio in the SELECT trial and lucinactant or poractant alfa in a 1:1 ratio in the STAR trial. In both trials, infants were randomly assigned by stratification on the basis of birth weight at each site. Clinicians who were caring for participating infants remained blinded to the surfactant assigned at randomization throughout their stay in the NICU and up to 1 year corrected age for prematurity (ie, chronologic age minus the number of weeks premature). Study protocols were approved by the institutional review boards of all participating centers, and signed informed consent was obtained for all participants. An independent data-safety board reviewed the study design and data for both trials. Both studies were conducted under the auspices of independent steering committees, which were chaired by the respective principal investigators.

From birth to 1 year corrected age, we recorded the occurrence of rehospitalization after discharge from the NICU, the number and type of respiratory illnesses that occurred after 36 weeks’ PMA (eg, wheezing, pneumonia, cough), and deaths. In addition, at the 1-year-corrected-age visit, weight, length, and head circumference were obtained and a physical examination, including a gross neurologic assessment, was performed. The neurologic examination assessed, at minimum, gross motor tone, reflex abnormalities, presence of unilateral or bilateral deafness, unilateral or bilateral blindness, and history of seizures that required treatment with anticonvulsant agents. The clinicians involved in the follow-up assessment phase of the studies also remained blind to the assigned surfactant treatment throughout the study.

**Statistical Analyses**

In this 1-year-corrected-age follow-up analysis, all randomly assigned infants were included on the basis of the intent-to-treat principle across both studies. In the STAR trial, the short-term results through 36 weeks’ PMA were previously reported on the basis of the per-protocol population (all infants who received any surfactant [\( N = 243 \)]), which is typical for noninferiority studies.13,14 For all survival comparisons, we used a prespecified imputation approach, with which loss to follow-up or withdrawal of consent was counted as a death. However, we also compared survival between groups using raw data without imputing loss to follow-up or consent withdrawal as death. The overall survival rate through 1 year corrected age for all randomly assigned infants within each study was estimated by using a standard Kaplan-Meier approach for long-term survival analysis. The Kaplan-Meier curves for lucinactant versus animal-derived surfactants in the combined analysis were estimated by using meta-analysis methods in which the overall Kaplan-Meier curves were constructed on the basis of the weighted average of the individual curves from the studies, weighted by study size.15 Also, when comparing lucinactant with animal-derived surfactants, we used meta-analysis methodology for analyzing data across studies with different sample sizes between and within these studies.1,15 This statistical method was chosen to compare these surfactants because simple pooling of data from the STAR and SELECT trials is not appropriate given that the randomization ratios in the 2 studies were unequal (lucinactant/colfosceril palmitate/beractant, randomization ratio: 2:2:1 [SELECT]; lucinactant/oractant alfa, randomization ratio: 1:1 [STAR]). Survival rates through 1 year corrected age for individual studies and the meta-analysis were compared by using the Wilcoxon test adjusting for study, birth weight strata, country, gender, and race. In addition to using the standard Kaplan-Meier approach for survival comparisons, we determined fixed-time-point estimates of mortality by imputing loss to follow-up as a death and also using raw data and compared them by using logistic regression adjusting for pooled center and birth weight stratum.

The incidence of rehospitalizations was analyzed by using logistic regression, and the total number of postdischarge rehospitalizations was compared by using analysis of variance. Data on respiratory morbidity
through 1 year corrected age were collected for those patients who were alive at 36 weeks’ PMA. The incidence of respiratory illnesses was compared by using logistic regression. The total number of respiratory illnesses through 1 year corrected age was compared by using analysis of variance. Neurologic outcomes at 1 year corrected age were only assessed for surviving infants in whom the data were captured; the data were compared by using the Cochran-Mantel-Haenszel test. All analyses described above were adjusted for pooled center and birth weight stratum. No missing data imputation was performed unless clearly specified.

RESULTS

Patient Population and Disposition

In the SELECT study, 527, 509, and 258 preterm infants between 24 and 32 weeks’ gestational age and between 600 and 1250 g birth weight were randomly assigned to receive lucinactant, colfosceril palmitate, and beractant, respectively. In the STAR study, 124 and 128 preterm infants in the same weight range were randomly assigned to receive lucinactant and poractant alfa, respectively. The upper limit for gestational age in the STAR trial was <29 weeks. Complete maternal and neonatal demographics for the STAR and SELECT study populations have been reported. A summary of the main demographic characteristics of both studies is given in Table 1. The study populations were fairly similar; small, nonsignificant differences in Apgar scores, birth weight, and prenatal steroid use were observed among infants enrolled in the 2 trials. In total, 651 infants were randomly assigned to lucinactant and 386 to animal-derived surfactants (beractant and poractant alfa).

In both trials and subsequent follow-up, 6 infants (0.9%) who received lucinactant and 8 (1.6%) and 6 (1.6%) infants who received colfosceril palmitate and an animal-derived surfactant, respectively, were lost to follow-up (Fig 1). Consent after randomization was withdrawn for 3 infants (0.5%) who received lucinactant and 2 (0.4%) and 4 (1.5%) infants who received colfosceril palmitate and an animal-derived surfactant, respectively. Per the prespecified rules (see “Methods”), these losses to follow-up were counted as deaths. Nonetheless, overall survival comparisons were also conducted only on infants for whom data were available (ie, without imputation for loss to follow-up as a death).

Overall Survival

The Kaplan-Meier survival curves through 1 year corrected age for infants in the SELECT study, the STAR study, and the combined group of infants who received lucinactant compared with those who received animal-derived surfactants (beractant and poractant alfa) are presented in Fig 2 (A, B, and C, respectively). Comparisons of mortality data for lucinactant versus the individual comparator surfactants at both 28 days’ and 36 weeks’ PMA in the SELECT and STAR trials have already been provided in their original publications. However, the findings of the combined analysis at 36 weeks’ PMA represent new data and are mentioned briefly below, because they provide a reference point for the 1-year-corrected-age data. There were no significant differences in the proportion of infants who were alive through 1 year corrected age comparing those given lucinactant with those who received colfosceril and beractant (Fig 2A). Fixed-time-point estimates of mortality at 1 year corrected age imputing loss to follow-up as a death were 28.1% for lucinactant, 31.0% for colfosceril, and 31.0% for beractant (OR: 0.81; 95% CI: 0.60–1.09 [lucinactant versus colfosceril]; OR: 0.84; 95% CI: 0.58–1.21 [lucinactant versus beractant]). In the STAR trial, significantly more infants who were treated with lucinactant were alive through 1 year corrected age compared with those who received poractant alfa (P = .04, Wilcoxon test; Fig 2B). Fixed-time-point estimates of mortality at 1 year corrected age imputing loss to follow-up as a death were 19.4% for lucinactant and 24.2% for poractant (OR: 0.64; 95% CI: 0.32–1.27). In the combined analysis, survival through 1 year corrected age of those infants who received lucinactant was higher compared with those who received animal-derived surfactants (P = .05, Wilcoxon test; Fig 2C). Fixed-time-point estimates using combined data from both trials imputing loss to follow-up as a death to compare lucinactant with the animal-derived surfactants demonstrated a lower mortality rate that favored infants who received lucinactant at 36 weeks’ PMA (20.1% vs 24.7%; P = .045; OR: 0.70; 95% CI: 0.50–0.99).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Summary of the Main Demographic Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
<td>SELECT Trial</td>
</tr>
<tr>
<td></td>
<td>Lucinactant</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>263 (49.9)</td>
</tr>
<tr>
<td>Birth weight, mean (SD), g</td>
<td>974 (183)</td>
</tr>
<tr>
<td>Gestational age, mean (SD), wk</td>
<td>28.2 (1.9)</td>
</tr>
<tr>
<td>Apgar score at 5 min, median (range)</td>
<td>7 (3–10)</td>
</tr>
<tr>
<td>Prenatal steroid use, n (%)</td>
<td>415 (79.2)</td>
</tr>
</tbody>
</table>
fixed-time-point estimates of mortality at 1 year corrected age imputing loss to follow-up as a death for lucinactant and animal-derived surfactants were 26.0% vs 29.4%, respectively (OR: 0.80; 95% CI: 0.58–1.09).

Because very few infants from both trials were either lost to follow-up or had their consent withdrawn, estimations of overall survival using raw data without imputation for loss to follow-up as a death introduced minor variations to the findings described above but did not modify the trend of the results. Using raw data without imputing loss to follow-up as a death, fixed-time-point estimates of mortality at 1 year corrected age were 26.6%, 29.1%, and 28.3%, respectively (OR: 0.83; 95% CI: 0.61–1.12 [lucinactant versus colfosceril]; OR: 

---

**FIGURE 1**
Enrollment and disposition of infants in the SELECT and STAR trials through 1 year corrected age. * Did not receive study medication.

**FIGURE 2**
Kaplan-Meier survival plots showing all-cause mortality through 1 year corrected age: A, data from the SELECT trial; B, data from the STAR trial; C, combined data from infants in both trials who received either lucinactant or animal-derived surfactants. All-cause mortality through 1 year corrected age favored lucinactant (Wilcoxon test) over poractant alfa (P = .04) and the animal-derived surfactants (P = .05).
Postdischarge Rehospitalization and Respiratory Morbidity

Although between one third and one half of the infants in both studies were rehospitalized during the first year, in general, readmission to the hospital occurred only once for most of them (Table 2). The incidence of postdischarge rehospitalization did not differ between surfactant groups in the SELECT trial (lucinactant: 43.4%; colfosceril: 43.4%; beractant: 50.8%) or the STAR trial (lucinactant: 34%; poractant alfa: 34.7%). Likewise, the number of postdischarge rehospitalizations through 1 year corrected age for all discharged infants did not differ between groups. Data on the number of respiratory illnesses (coughing, wheezing, and pneumonia) through 1 year corrected age were collected for infants from both studies who were alive at 36 weeks’ PMA. There were no significant differences between the groups (Table 2).

Neurologic Assessment

Neurologic evaluations at 1 year corrected age were performed for most surviving infants from the 2 studies. In the SELECT trial, 731 (81%) of the 908 infants alive at 1 year corrected age had neurologic evaluations (lucinactant: 306; colfosceril palmitate: 279; beractant: 146). In the STAR trial, neurologic evaluations were performed in 190 (96%) of the 197 infants alive at 1 year corrected age (lucinactant: 98; poractant alfa: 92; Table 3).

Among survivors examined from the SELECT and STAR trials, there were generally no significant differences between groups with very few exceptions (Table 3). There were fewer muscle-tone abnormalities in the infants in the lucinactant group compared with those in the colfosceril and beractant groups and also a lower incidence of gross motor delay compared with the colfosceril group. There were no major differences in neurologic outcomes at 1 year corrected age when comparing infants who received lucinactant with those who received animal-derived surfactants (Table 4) despite the greater number of survivors in the lucinactant group.

DISCUSSION

Surfactant therapy is associated with an absolute reduction of ~5% to 7% in neonatal mortality among preterm infants compared with those who receive placebo.16,17 This reduction in mortality rate translates into ~1 life saved for every 14 to 20 infants who receive surfactant. The other major benefit consistently shown in trials that have compared surfactant administration with placebo is a marked decrease in the occurrence of air leaks. On the basis of these findings, surfactant therapy has become widely used for the prevention and treatment of RDS.

There have been many randomized trials that compared the 2 major classes of surfactants, namely synthetic preparations devoid of SPs and animal-derived surfactants that contain variable amounts of SP-B and SP-C, although most have used a treatment rather than a prophylactic approach. Soil and Blanco1 conducted a systematic review of 11 trials and compared animal-derived surfactants versus synthetic surfactants. Ten of the trials included in that review compared colfosceril with beractant (7 trials),18–24 calfactant (2 trials),25,26 or poractant alfa (1 trial),27 whereas only 1 trial compared pumactant to poractant.28 These authors concluded that both types of surfactants are effective in the treatment and prevention of RDS. They also concluded that when taken together, use of animal-derived surfactants resulted in fewer deaths, greater early improvement in the requirement for ventilatory support, and a lower overall incidence of pneumothorax than synthetic products that

TABLE 2

Postdischarge Rehospitalizations and Respiratory Illnesses From 36 Weeks’ PMA to 1 Year Corrected Age

<table>
<thead>
<tr>
<th></th>
<th>SELECT Trial</th>
<th>STAR Trial</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lucinactant</td>
<td>Colfosceril Palmitate</td>
</tr>
<tr>
<td>No. of postdischarge rehospitalizations n/N (%)</td>
<td>173/399 (43.4)</td>
<td>159/366 (43.4)</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>0.8 (1.4)</td>
<td>1.0 (1.7)</td>
</tr>
<tr>
<td>Median (range)</td>
<td>0.0 (0.0–8.0)</td>
<td>0.0 (0.0–8.0)</td>
</tr>
<tr>
<td>No. of respiratory illnesses n/N (%)</td>
<td>281/406 (69.2)</td>
<td>263/375 (70.1)</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>1.9 (2.3)</td>
<td>2.1 (2.1)</td>
</tr>
<tr>
<td>Median (range)</td>
<td>1.0 (0.0–26.0)</td>
<td>2.0 (0.0–13.0)</td>
</tr>
</tbody>
</table>

* N = all discharged infants through 1 year corrected age.
* N = all infants alive as of 36 weeks’ PMA with available data.
TABLE 3 Abnormal Gross Neurologic Findings: All Randomly Assigned Evaluated Survivors in Each Trial

<table>
<thead>
<tr>
<th>Neurologic Finding</th>
<th>SELECT Trial</th>
<th>STAR Trial</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lucinactant (N = 306)</td>
<td>Colfosceril Palmitate (N = 279)</td>
</tr>
<tr>
<td>Gross tone or reflex abnormality, n (%)</td>
<td>30 (9.8)</td>
<td>45 (16.1)*</td>
</tr>
<tr>
<td>Gross motor delay, n (%)</td>
<td>30 (9.8)</td>
<td>42 (15.1)*</td>
</tr>
<tr>
<td>Unilateral or bilateral deafness, n (%)</td>
<td>10 (3.3)</td>
<td>7 (2.5)</td>
</tr>
<tr>
<td>Unilateral or bilateral blindness (or registered as blind), n (%)</td>
<td>8 (2.6)</td>
<td>10 (3.6)</td>
</tr>
<tr>
<td>Seizures requiring anticonvulsant agents, n (%)</td>
<td>7 (2.3)</td>
<td>7 (2.5)</td>
</tr>
</tbody>
</table>

a P ≤ .05 versus lucinactant. No other comparisons were significant.

TABLE 4 Pooled Analysis of Abnormal Gross Neurologic Findings: All Randomly Assigned Evaluated Survivors

<table>
<thead>
<tr>
<th>Neurologic Finding</th>
<th>Lucinactant (N = 404)</th>
<th>Animal-Derived Surfactants (N = 238)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gross tone or reflex abnormality, %</td>
<td>9.4</td>
<td>15.3*</td>
</tr>
<tr>
<td>Unilateral or bilateral deafness, %</td>
<td>2.5</td>
<td>0.5</td>
</tr>
<tr>
<td>Unilateral or bilateral blindness (or registered as blind), %</td>
<td>2.2</td>
<td>1.6</td>
</tr>
<tr>
<td>Seizures requiring anticonvulsant agents, %</td>
<td>2.0</td>
<td>2.9</td>
</tr>
</tbody>
</table>

No comparisons were significant: *P = .05 versus lucinactant.

contain only phospholipids. However, none of the studies that compared beractant to colfosceril, including large, well-conducted trials, have shown a significant difference in overall mortality rate favoring either surfactant, either singly or in combination. Furthermore, in the only trial that compared poractant alfa and colfosceril, the overall mortality rate was higher in the poractant group (20%) than in the colfosceril group (13%), but this difference was not significant, probably because of the relatively small sample size of the study (N = 228).

No advantages in terms of reduction in the incidence of BPD were described in this systematic review or in any of the individual trials included in it. In addition, a small but significant increase in the incidence of intraventricular hemorrhage among infants who were treated with animal-derived surfactants was reported.

We recently reported the results of 2 multicenter, randomized, double-blind trials that compared the new-generation synthetic surfactant lucinactant, which contains a peptide that mimics the main function of human SP-B, with other synthetic or animal-derived surfactants for the prevention of RDS.11,12 In the largest of these trials, the SELECT study, lucinactant was shown to decrease the incidence of RDS, RDS-related mortality, and BPD compared with colfosceril, but no reduction in overall mortality rate compared with this surfactant was observed.11 These findings supported the hypothesis that the addition of a peptide that mimics the main action of SP-B to surfactant phospholipids improves short-term clinical outcomes compared with using a surfactant that contains only phospholipids. In this study, a reference arm of infants who were randomly assigned to receive beractant was included. Lucinactant reduced RDS-related deaths and the overall mortality rate at 36 weeks’ corrected age compared with beractant, but there was no difference in BPD. It is notable that this is the only study to date that has compared prophylactic administration of colfosceril and beractant. Although this was not the comparison of primary interest in the SELECT trial, the findings of a lower incidence of RDS at 24 hours and more rapid weaning with beractant than with colfosceril parallel those of previous randomized comparison trials of these 2 types of surfactants for the treatment of established RDS.18,19 Similar to all previous trials that compared colfosceril and beractant, the SELECT trial did not demonstrate superiority of beractant over colfosceril in terms of BPD or overall mortality rate. The smaller of the 2 randomized trials of lucinactant, the STAR study, compared this surfactant with poractant alfa.12 This trial is the second largest study published to date to compare poractant alfa to another surfactant. In this trial, the phospholipid doses of both surfactants were similar (175 mg/kg) but were also higher than in other comparison trials of surfactants (for doses and administration in both trials, see the original publications11,12). Furthermore, both poractant alfa and lucinactant contain more SP-B (or its equivalent as sinapultide) than beractant. The primary outcome of being alive without BPD at 28 days was observed to be more frequent in the lucinactant group (37.8%) than in the poractant alfa group (33.1%), but without statistical significance. No differences in other secondary outcomes between groups were identified.

In these 2 lucinactant trials there were far fewer differences in study design and the populations studied than between those studies included in the systematic review by Soll and Blanco.1 Furthermore, most outcomes evaluated in the lucinactant trials were within the ranges reported in those studies included in the review by Soll and Blanco and data from the Vermont Oxford Network.29 In view of the relatively similar design of the lucinactant trials and considering that lucinactant is a different class of surfactant than previous synthetic and animal-derived preparations, we sought to compare overall survival between infants who received lucinact-
tant versus those who received the other types of surfactants, not only within each trial but also by using combined data from both trials. We elected to present overall survival data by using the standard Kaplan-Meier approach, because it allows for comparison of survival between groups through the entire observation period (up to 1 year corrected age) and also because this methodology was used to report survival through 36 weeks’ PMA in both of the original publications of the lucinactant trials.11,12 Using a similar approach to report the 1-year survival curves should facilitate comparison with previous data. Kaplan-Meier survival estimates do not falsely amplify treatment differences with respect to survival. Rather, they generally reveal unbiased estimates of survival rates for each treatment, hence yielding unbiased estimates of treatment differences. Nonetheless, because most of the surfactant benefit in mortality occurs in the neonatal period, in our survival analysis comparing treatments through 1 year corrected age we used the Wilcoxon test, which emphasizes earlier treatment differences. Using this approach for analysis, lucinactant administration resulted in better survival through 1 year corrected age than poractant in the STAR trial but no difference with colfosceril or beractant in the SELECT trial. In the combined analysis there was better survival at 1 year corrected age for lucinactant versus the animal-derived surfactants (beractant and poractant), which was of borderline statistical significance but potentially of clinical importance.

We also calculated fixed estimates of mortality at defined time points to allow for crude estimations of the relative magnitude of change and its CIs by using both the prespecified worst-case scenario approach (imputing loss to follow-up as a death) and raw data. The impact of imputations on treatment differences is uncertain and depends on variations in the rate of censoring among treatment groups. When there are premature withdrawals from a study before the end of follow-up, a raw incidence estimate may be biased (ie, counting premature withdrawals as death will generally overestimate the event rate, whereas counting premature withdrawals as survival will generally underestimate the event rate). Not unexpectedly, results of the fixed-time-point estimates of mortality depended on whether imputation was used. Regardless of the methodology used, mortality estimates for infants who received lucinactant were either comparable or significantly lower than those observed for the other surfactants. Unfortunately, only short-term mortality data from the previously published surfactant-comparison trials are available, none of which evaluated survival at 1 year corrected age.

Several randomized trials that compared different animal-derived surfactants for prevention and treatment of RDS have been conducted in the past decade.30–34 Some of these surfactants contain more SP-B (calfactant, poractant) or have used higher doses of phospholipids than when beractant is administered.2,3 Therefore, it is not surprising that a faster improvement in oxygenation versus beractant has been observed in some of them.30,33,34 However, no differences in overall mortality rate have been reported in the 2 large RDS-prevention or -treatment trials that compared calfactant and beractant.30,31 Several relatively small trials have compared poractant with beractant only for treatment of RDS.35 These studies administered poractant using either a higher initial dose of phospholipids (200 mg/kg) or a similar amount (100 mg/kg) compared with beractant. In a preliminary meta-analysis of these trials, Halliday35 suggested that administration of poractant resulted in a lower neonatal mortality rate than beractant primarily when the higher initial dose was given, because no difference in mortality rate versus beractant was found in those trials that administered 100 mg/kg poractant initially. Using this lower initial dose for treatment of RDS, poractant administration resulted in a higher mortality rate (20%) than colfosceril (13%), although this difference did not achieve statistical significance, probably because of the study’s sample size.27 At present, it is impossible to differentiate whether any potential benefits of the higher initial dose of poractant are a result of administration of more phospholipids, a higher amount of SP-B, a higher volume of drug (2.5 vs 1.25 mL/kg), which may improve lung distribution, or other alternative explanations. Nonetheless, Halliday concluded that for infants with moderate-to-severe RDS, the larger dose of poractant is more effective, but for prophylaxis a lower dose may be appropriate; however, this hypothesis has yet to be tested prospectively in a clinical trial.31 In keeping with this notion and because, to our knowledge, no randomized comparison of these 2 surfactants for prevention of RDS using any dose has ever been conducted, we combined data from infants who received beractant and poractant in both of the lucinactant trials for analysis. Furthermore, we used appropriate methodology for analyzing data across studies with different sample sizes and randomization schemes.1,15

To our knowledge, previous surfactant-comparison trials have not reported outcomes that were prospectively collected beyond the neonatal period. This may limit the ability to draw conclusions related to the impact of surfactants in long-term survival and other morbidity frequently observed in preterm infants after discharge from neonatal intensive care. Nonetheless, published follow-up data from studies that compared various surfactants with placebo demonstrate that the improved survival observed resulting from surfactant treatment is not associated with increased subsequent morbidity.36 Because lucinactant is a new-generation surfactant, we sought to determine postdischarge morbidity and outcome up to 1 year corrected age of infants who received this as well as the other surfactants used in these 2 trials. We successfully collected this information in nearly all of
the participants or survivors from both trials. Our data indicate that rehospitalization of preterm infants after discharge from neonatal intensive care remains a frequent event and confirm the findings from previous follow-up studies of surfactant therapy.37 When we examined the incidences of respiratory morbidity and rehospitalization in each trial, no differences were detected between infants given the various surfactant preparations, including lucinactant. Similarly, neurologic assessment of these infants showed essentially no differences except for a lower occurrence of muscle-tone abnormalities and gross motor delay favoring the lucinactant group, even with a higher number of survivors in the lucinactant group. Although our data on neurologic evaluations have limitations, most infants from both trials were assessed at 1 year corrected age. In fact, the proportion of infants we were able to follow was similar or higher than many of the previous follow-up studies of infants treated with surfactant.38,39 An additional strength of these evaluations is that they were conducted by physicians who were unaware of group assignment.

**CONCLUSIONS**

The findings of this 1-year follow-up study, as well as data from the original reports of the lucinactant trials, strongly suggest that administration of lucinactant to infants at risk for RDS results in neonatal survival that is at least comparable with, if not superior to, that of infants given the animal-derived surfactants beractant and poractant alfa. Furthermore, these data demonstrate that there is no difference in morbidity through 1 year corrected age in infants given lucinactant versus other animal-derived surfactants despite the proportionally higher number of survivors. These findings strongly support the long-term safety of using lucinactant for the prevention of neonatal RDS.

**ACKNOWLEDGMENTS**

This study was funded by Discovery Laboratories, Inc.

We acknowledge respective members of the SELECT and STAR Steering Committees, members of the Data Safety Monitoring Board, participating investigators and their staff, and Discovery Laboratories personnel, who monitored the study. Thomson Scientific Connections provided assistance with manuscript styling and graphics preparation.

**REFERENCES**

22. Modanlou HD, Beharry K, Padilla G, Norris K, Salvati S, Aranda JV. Comparative efficacy of Exosurf and Survanta sur-


Adherence to Antiretroviral Therapy for Pediatric HIV Infection: A Qualitative Systematic Review With Recommendations for Research and Clinical Management

Jane M. Simoni, PhD, Arianna Montgomery, MPH, Erin Martin, MPH, Michelle New, PhD, Penelope A. Demas, PhD, Sohail Rana, MD, FAAP

Department of Psychology, University of Washington, Seattle, Washington; Center for HIV/AIDS/Viral Hepatitis/STD, and TB Prevention, Division of HIV/AIDS Prevention, Centers for Disease Control and Prevention, Atlanta, Georgia; Departments of Special Immunology and Psychology, Children’s National Medical Center, Washington, DC; AIDS Research Program, Montefiore Medical Center, Bronx, New York; Department of Pediatrics and Child Health, Howard University, Washington, DC

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

ABSTRACT

Although nonadherence to prescribed therapies is widespread, it is particularly problematic with highly active antiretroviral therapy for HIV infection. This review of >50 studies in the area of pediatric HIV infection revealed varying methods for assessing antiretroviral adherence with a wide range of estimates of adherence. Correlates of adherence could be grouped as those relating to the medication, the patient, and the caregiver/family, with many conflicting findings and a lack of theory guiding the research. Only 8 studies, mainly small feasibility or pilot investigations, evaluated highly active antiretroviral therapy adherence interventions in pediatric populations. We conclude with specific recommendations for assessment and clinical management of adherence and discuss directions for future research in this area.
Adherence to therapy, or the extent to which a patient’s behaviors coincide with medical advice, mutually negotiated between the health professional and the patient, is a universal challenge with all illnesses and in all age groups. Each year in the United States, 500,000 physicians write 1.8 billion prescriptions involving 55,000 pharmacies. However, many of these prescribed medications are never taken, with rates of nonadherence ranging from 15% to 93%. Among persons with chronic illnesses, nonadherence is especially problematic, occurring in up to 82% of cases. The effects of nonadherence range from individual disability (e.g., unrelied pain) to global threat (e.g., development of treatment-resistant bacteria or viruses). The yearly monetary costs of nonadherence exceed $100 billion.

Adherence is particularly critical with highly active antiretroviral therapy (HAART) in the treatment of pediatric HIV infection. The data on HAART for pediatric HIV infection, although scarcer than for adults, suggest that medication adherence is a strong predictor of therapeutic impact. For example, Wiener et al. observed that among children with an HIV-1 RNA viral load (VL) <10,000, 75% had taken 100% of their medication doses in the previous week, whereas among those with a VL of ≥10,000, only 36% reported taking all of their medication.

Despite the benefits of HAART in treating pediatric HIV infection and the adverse consequences of nonadherence, adherence is reportedly suboptimal among children. It is likely thwarted by multiple barriers and complicated because, unlike with many other chronic illnesses, most children who are born with HIV in the United States are ethnic/racial minorities who live in chronic poverty with limited resources and face discrimination, family disruption, substance abuse, and the stressors of life in the inner city. In addition, stigma is greater for HIV/AIDS than other chronic illnesses, which often leads caregivers to conceal the child’s diagnosis and treatment from others and from the child as well. The energy spent on maintaining this secrecy can strain the family considerably. The practical barriers and psychological burden further complicate the administration of medications, which itself often serves as a painful reminder of the disease.

The global epidemic of HIV continues, with an increasing impact on children. By the end of 2004, an estimated 2.2 million children under 15 years of age worldwide were living with HIV; in that year alone, 640,000 were newly infected. Although recent meta-analyses suggest that efforts to improve adherence to antiretroviral therapy among adults can be effective, much work remains to be done with children.

In an effort to guide clinicians and researchers with respect to pediatric HAART adherence, we conducted a qualitative systematic review of the research literature. We present (1) estimates of adherence and examine how they vary according to assessment method, (2) correlates of adherence (including barriers to adherence and reasons for nonadherence), and (3) strategies to improve adherence, with particular attention to those that have been evaluated empirically.

**METHODS**

A thorough and systematic search for published journal articles, abstracts, books, and ongoing studies on pediatric adherence to antiretroviral therapy was conducted via PubMed, PsycINFO, Medline, AIDSLINE, and the Computer Retrieval of Information on Scientific Projects, an online database of ongoing federally funded biomedical research projects. We searched each database for work from January 1996 to December 2005 that included the terms “HIV/AIDS,” “pediatric,” and “compliance/adherence” (equivalents for these terms also were used in the searches). In addition, we examined the reference lists of relevant articles and solicited information from colleagues. A team of 3 reviewers narrowed down the ~350 citations identified in the search to the ~50 that we cite in this review, mainly on the basis of their offering original data on the topic. We conducted a qualitative review of this literature and summarized nonnumerically the major findings.

**RESULTS**

**Estimates of Adherence According to Assessment Method**

As shown in Table 1, we found 32 studies that reported data that estimated antiretroviral adherence among pediatric populations. They were published between 1999 and 2005, with 69% of the studies based in the United States. Sample sizes ranged from 10 to 262 (median = 46), and participants were from 3 months to 24 years of age. Most participants were infected perinatally, but behaviorally infected individuals and those infected by contaminated blood products also were included in some studies.

The authors used most of the known methods for capturing antiretroviral medication adherence, including both direct and indirect methods. Direct methods such as biological assays of an active drug, its metabolite, or other markers in the bodily fluids confirm active drug ingestion. Indirect methods, which do not measure the presence of the drug in the individual, include self-report; caregiver report; clinician assessment; medical chart review; clinic attendance; pill count; pharmacy refill records; electronic drug monitoring (EDM); behavioral observation in the form of directly observed therapy (DOT) (sometimes during hospitalization); resistance testing; and therapeutic impact such as VL, CD4 lymphocyte count, Centers for Disease Control and Prevention–defined stage of disease progression, and mortality. Combinations of these methods to produce one adherence estimate also were used.
<table>
<thead>
<tr>
<th>Assessment Interval, d</th>
<th>Study Description</th>
<th>N</th>
<th>Age</th>
<th>Route of Infection</th>
<th>Location</th>
<th>Ref No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>Caregiver report of 100% adherence, % adherent</td>
<td>3</td>
<td>26</td>
<td>21 mo to 12.5 y</td>
<td>IP</td>
<td>Baltimore, MD</td>
</tr>
<tr>
<td>97</td>
<td></td>
<td>7</td>
<td>42</td>
<td>4 mo to 18 y</td>
<td>NR</td>
<td>New York, NY</td>
</tr>
<tr>
<td>83</td>
<td></td>
<td>1</td>
<td>90</td>
<td>15 mo to 14 y</td>
<td>NR</td>
<td>Boston, MA</td>
</tr>
<tr>
<td>79.5</td>
<td></td>
<td>3</td>
<td>44</td>
<td>1 to 17 y</td>
<td>IP</td>
<td>Genova, Italy</td>
</tr>
<tr>
<td>76</td>
<td></td>
<td>3</td>
<td>37</td>
<td>8 y (median)</td>
<td>IP</td>
<td>Genova, Italy</td>
</tr>
<tr>
<td>74</td>
<td></td>
<td>7</td>
<td>262</td>
<td>3 mo to 16 y</td>
<td>NR</td>
<td>Europe and Brazil</td>
</tr>
<tr>
<td>73</td>
<td></td>
<td>3</td>
<td>11</td>
<td>4 mo to 19 y</td>
<td>NR</td>
<td>Belgium</td>
</tr>
<tr>
<td>68.2</td>
<td></td>
<td></td>
<td>30–60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>67</td>
<td></td>
<td>21</td>
<td>18</td>
<td>4 to 15.5 y</td>
<td>IP</td>
<td>Sydney, Australia</td>
</tr>
<tr>
<td>60</td>
<td></td>
<td>30</td>
<td>75</td>
<td>3 to 13 y</td>
<td>IP</td>
<td>New York, NY</td>
</tr>
<tr>
<td>57</td>
<td></td>
<td>7</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td></td>
<td></td>
<td>34</td>
<td>18 mo to 20 y</td>
<td>NR</td>
<td>South Carolina</td>
</tr>
<tr>
<td>89</td>
<td>Caregiver report of ≥95% adherence, % adherent</td>
<td>21</td>
<td>129</td>
<td>4 mo to 17 y</td>
<td>NR</td>
<td>Multisite, Italy</td>
</tr>
<tr>
<td>84</td>
<td></td>
<td>4</td>
<td>30</td>
<td>129</td>
<td>NR</td>
<td>Multisite, Italy</td>
</tr>
<tr>
<td>96</td>
<td>Caregiver report, mean adherence, %</td>
<td>1</td>
<td>40</td>
<td>2 to 17 y</td>
<td>IP</td>
<td>Detroit, MI, and New York, NY</td>
</tr>
<tr>
<td>94.3</td>
<td></td>
<td>3</td>
<td>30</td>
<td>0.9 to 11.5 y</td>
<td>NR</td>
<td>Memphis, TN</td>
</tr>
<tr>
<td>88.4</td>
<td></td>
<td></td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>58</td>
<td>Self-report of 100%, % adherent</td>
<td>1</td>
<td>42</td>
<td>16 to 24 y</td>
<td>NR</td>
<td>Chicago, IL</td>
</tr>
<tr>
<td>44</td>
<td></td>
<td>30</td>
<td>161</td>
<td>12 to 19 y</td>
<td>IB</td>
<td>Multisite, US</td>
</tr>
<tr>
<td>34</td>
<td></td>
<td>7</td>
<td>7</td>
<td>35</td>
<td>11 to 21 y</td>
<td>40% IP, 60% BP</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>14</td>
<td>7</td>
<td>35</td>
<td>35</td>
<td>40% IP, 60% BP</td>
</tr>
<tr>
<td>44</td>
<td>Self-report of ≥95% adherence, % adherent</td>
<td>14</td>
<td>88</td>
<td>13 to 22 y</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>69.8</td>
<td></td>
<td>30</td>
<td>231</td>
<td>15 to 22 y</td>
<td>IB</td>
<td>Multisite, US</td>
</tr>
<tr>
<td>61</td>
<td></td>
<td>90</td>
<td>31</td>
<td>13 to 24 y</td>
<td>IB</td>
<td>Multisite, US</td>
</tr>
<tr>
<td>58</td>
<td></td>
<td></td>
<td>1</td>
<td>31</td>
<td>21 y</td>
<td></td>
</tr>
<tr>
<td>97</td>
<td>Self-report, mean adherence, %</td>
<td>30</td>
<td>88</td>
<td>13 to 22 y</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>93</td>
<td></td>
<td></td>
<td>88</td>
<td>13 to 22 y</td>
<td>NR</td>
<td>Cape Town, South Africa</td>
</tr>
<tr>
<td>Pharmacy refill</td>
<td>100% filled</td>
<td>90</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>92% to 100% filled</td>
<td>180</td>
<td>92% to 100% adherent</td>
<td>365</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49% to ≥90% adherent</td>
<td>90</td>
<td>49% to ≥90% adherent</td>
<td>365</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>58% filled ≥75%</td>
<td>180</td>
<td>58% filled ≥75%</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pill count</td>
<td>&quot;Most&quot; had 85%</td>
<td>30</td>
<td>80</td>
<td>NR</td>
<td>NR</td>
<td>Cape Town, South Africa</td>
</tr>
<tr>
<td>89.9% adherent (mean)</td>
<td>90</td>
<td>89.9% adherent (mean)</td>
<td>38</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Care provider report</td>
<td>82.5% adherent (mean)</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>82% adherent (mean)</td>
<td>30</td>
<td>82% adherent (mean)</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>97% reported 100% adherence</td>
<td>NR</td>
<td>97% reported 100% adherence</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinic attendance</td>
<td>88% kept 100%</td>
<td>90</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65% kept 100%</td>
<td>180</td>
<td>65% kept 100%</td>
<td>34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21% kept 100%</td>
<td>365</td>
<td>21% kept 100%</td>
<td>33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Therapeutic drug monitoring, % adherent</td>
<td>66</td>
<td>50</td>
<td>3.8 to 16.8 y</td>
<td>NR</td>
<td>PACTG cohort, US</td>
<td>86</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>40</td>
<td>3 to 18 y</td>
<td>NR</td>
<td>Rotterdam, Netherlands</td>
<td>35</td>
</tr>
<tr>
<td>54–100</td>
<td></td>
<td></td>
<td>40</td>
<td>3 mo to 18 y</td>
<td>NR</td>
<td>Rotterdam, Netherlands</td>
</tr>
</tbody>
</table>

TABLE 1 Estimates of Pediatric Antiretroviral Adherence According to Assessment Method
Notably, evidence for drug resistance, increasing VL, and decreasing CD4 lymphocyte count are commonly considered signs of nonadherence, but they are not well represented as assessment strategies in research studies of pediatric HAART adherence. In particular, VL is often assessed clinically, but in the research literature it is most commonly used to corroborate adherence reports or as an indicator of treatment success rather than as an estimate of adherence itself.

Caregiver and patient reports were the sole or main source of adherence data for most reports. Few studies used purportedly more-objective methods such as pill counts and EDM. Some studies used more than 1 method to collect adherence data but reported just 1 overall adherence estimate; this was done in 3 studies in which older children provided self-report data while caregivers provided data for the younger children and infants.13,28,29 Other studies combined self-report with clinical attendance data9 or provider reports with pill counts.30

Adherence estimates varied greatly, precluding a meaningful summary of the extent of nonadherence. As seen in Table 1, results from the 13 studies that used caregiver reports indicated that 34% to 100% of caregivers reported 100% adherence, and 84% to 89% reported 95% adherence; reports of mean adherence according to caregivers ranged from 88.4% to 96%. Self-reports, which were used in 8 studies overall, indicated that 20% to 58% of patients reported 100% adherence, 44% reported 95% adherence, and 58% to 70% reported 90% adherence; self-reported mean adherence ranged from 93% to 97%. In the 5 studies with pharmacy refill data, assessment intervals ranged broadly (from 90 to 365 days), and widely different definitions of adherence were used (eg, ≥75% filled, ≥90% filled, median filled). Thus, it is difficult to provide a single summary estimate of adherence on the basis of pharmacy refill data. For example, 1 study reported that 100% of prescriptions were refilled,31 and another reported that only 58% of participants filled >75% of their prescriptions.32 Fairly high adherence was estimated by the 2 studies by using pill counts (in which mean adherence was ~87%) and in the 3 studies using care providers (in which adherence averaged 82%–100%). In the 3 studies that assessed adherence by clinic attendance, estimates varied almost linearly according to the assessment interval, with lower levels for longer intervals. For example, perfect clinic attendance was achieved by 21% of the patients in the previous 365 days,33 65% in the previous 180 days,34 and 88% of patients in the previous 90 days.31,33 As seen in Table 1, the 3 studies that used therapeutic drug monitoring, the 2 that relied on medical chart review, and the 2 with EDM generally produced lower estimates of adherence than did studies using the other methods.

The range in estimates of adherence may reflect the heterogeneity of the samples. For example, age varied widely both within and across samples, and medication-taking among 4-month-olds is obviously a very different phenomenon than among adolescents. Another possible reason for the variation is that study participants were on varying antiretroviral regimens, and adherence can differ according to medication as a result of variations in such factors as palatability or tolerability.35 For example, in 1 study, full adherence varied from a low of 59% to a high of 100% depending on the regimen.29

<table>
<thead>
<tr>
<th>Results Assessment Interval, d</th>
<th>Study Description</th>
<th>N</th>
<th>Age Route of Infection Location</th>
<th>Ref No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical chart review, % adherent</td>
<td></td>
<td>54</td>
<td>NR</td>
<td>70 NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>28</td>
<td>NR</td>
<td>18 13 to 21 y</td>
</tr>
<tr>
<td>EDM</td>
<td></td>
<td>180</td>
<td>NR</td>
<td>90</td>
</tr>
<tr>
<td>Self-report and caregiver report of 100% adherence (combined), % adherent</td>
<td></td>
<td>70</td>
<td>3</td>
<td>125 4 mo to 17 y</td>
</tr>
<tr>
<td></td>
<td></td>
<td>67</td>
<td>30</td>
<td>112 4.3 to 13 y</td>
</tr>
<tr>
<td></td>
<td></td>
<td>54a</td>
<td>2</td>
<td>48 7.3 to 13.7 y</td>
</tr>
<tr>
<td>Self-report and clinic attendance (combined)</td>
<td></td>
<td>168</td>
<td>27% “good” 180</td>
<td>120 8 to 22 y</td>
</tr>
<tr>
<td>Pharmacy refill and pill count (combined)</td>
<td></td>
<td>19% to 95% were “adherent” 70% (mean)</td>
<td>21</td>
<td>3 mo to 15 y</td>
</tr>
</tbody>
</table>

NR indicates not reported in the original article; IP, infected perinatally; IB, infected behaviorally; BP, infected by blood products; PACTG, Pediatric AIDS Clinical Trials Group.

a In 54% of all patient/caregivers dyads, both reported 100% adherence.
Differences in study methodology also may have contributed to the wide range of adherence estimates. How adherence was assessed (eg, pill counts or self-report); how optimal adherence was defined (eg, cut points of 90%, 95%, or 100%); and how findings were reported in each study (eg, as a range, a median percentage for the whole sample, or the percentage of participants who achieved a predetermined benchmark) varied considerably. Furthermore, some studies counted only doses taken on time as completed.\textsuperscript{10} Assessment intervals (when reported) varied from 1 day to 1 year and were associated with adherence at least as assessed by clinic attendance.

The studies that used multiple adherence-assessment strategies in the same sample (ie, triangulation) allowed for a more useful comparison of adherence estimates because the differences in samples were controlled. The 6 studies in our review that used multiple methods to capture adherence, which were all based in the United States, are described below.

With 5 separate methods, Farley et al\textsuperscript{31} assessed adherence among 26 children aged 21 months to 12.5 years. Median adherence was found to be higher according to caregiver report (100%) and clinic attendance (100%) than pharmacy refill (92%) and EDM (81.4%). Data on provider assessment were not reported.

Among 42 patients aged 4 months to 18 years, Byrne et al\textsuperscript{31} used 4 methods of assessing adherence. They noted full adherence in the 7 days before assessment for 97% of the sample according to caregiver report (mean: 88.4%). Over 90 days, they noted full adherence for 88% according to clinic attendance and 100% according to pharmacy refill. Median adherence according to the providers was 82.5% (assessment interval was not reported).

Three other studies also concluded that caregivers generally overestimated adherence compared with other methods.\textsuperscript{36–38} With 73 children under the age of 13 years, Mellins et al\textsuperscript{36} reported that full adherence in the 30 days before assessment was 60% according to caregiver report versus 44% according to self-report. In the Naar-King et al\textsuperscript{37} sample of 40 children aged 2 to 17 years, mean adherence the day before the interview was 96% according to caregivers and 93% according to self-report. In 30 children under 12 years of age, Steele et al\textsuperscript{38} found that caregiver reports of adherence in the 3 days before assessment averaged 94.3%; over the previous 90 days, adherence according to pill counts averaged 89.9% and that according to EDM averaged 44%.

Wiener et al\textsuperscript{10} used a multisite sample of 35 individuals aged 11 to 21 years to examine adherence with provider ratings, self-report interviews, and daily telephone diaries. They found that providers (ie, clinical nurses) reported uniformly high adherence (all but 1 estimated it was ≥90%). Over the week before the assessment, self-reported full adherence on the basis of doses taken on time (12%) was less than for doses taken at all (32%); both were less than the 44% adherence estimated based on daily telephone diaries that did not directly query about medication-taking.

**Factors Related to Adherence**

Researchers commonly investigate factors related to adherence as answers to questions such as: What are the challenges or barriers to adherence? What are reported reasons for nonadherence? Which variables predict nonadherence, or what might facilitate optimal adherence? Lists of factors presented are more likely to be derived from clinical experience or surveys in selected areas such as “mental health and coping”\textsuperscript{39} than from theory-driven research or studies in which the associations were examined statistically.\textsuperscript{19,31,40–42}

Pontali\textsuperscript{41} grouped factors that were capable of influencing adherence into those related to the medication, health care system, and patient and family/caregiver. They stressed the role of family disruption and the characteristics of the caregiver such as his or her relationship to the child, level of anxiety and depression, education, substance use, attitudes toward treatment, and his or her own HIV infection status.

Proposed medication or treatment-related factors that likely complicate pediatric adherence include the indefinite duration of treatment; multiple and precise dosing times; multiple medications; high pill burden; complex dietary considerations; storage requirements; low palatability (ie, bitter-tasting liquids and gritty, sticky powders); large pills; significant short-term and long-term adverse effects (eg, nausea, rashes, hypersensitivity reactions, lipodystrophy, and anemia); and long-term toxicities.\textsuperscript{43,44} Studies have demonstrated better adherence with a twice-per-day (vs 3-times-per-day) nelfinavir regimen, shorter length of time since treatment initiation, and nelfinavir rather than indinavir but not with pill burden or drug toxicity.\textsuperscript{18,29,35,36}

Pontali\textsuperscript{41} cited such factors as the availability and cost of the medications, accessibility of treatment, and health care providers’ experience and relationship to the patient as health care system variables that are possibly associated with adherence. We found no studies that specifically evaluated the association between such variables and pediatric antiretroviral adherence.

In pediatric disease, developmental challenges, especially during adolescence, are the major patient factors that complicate HAART adherence. Marhefka et al\textsuperscript{41} wrote that adolescents have the same difficulties as adults with fitting a complex regimen into their life, but they have less autonomy, privacy, and mobility than adults. A qualitative study of 6 HIV-infected teenagers aged 16 to 24 years in Rhode Island found that many of the major factors that negatively influenced adherence were the same as those in adults. These factors included fear of social stigma related to HIV disclosure, complexity
of the regimen, adverse effects, forgetfulness, depression, absence of symptomatology, lack of general knowledge about HIV, and poor provider-patient relationship. However, the authors observed other factors that were unique to the developmental stage of the teenagers: familial overinvolvement and the youth’s focus on gaining rewards in the present versus planning for the future. Similarly, in a study of 29 HIV-positive French adolescents, participants reported that in addition to loneliness and depression, secrecy and silence about their HIV diagnosis and treatment were major barriers to their sense of freedom, with nonadherence becoming a way to express their need for autonomy. Given these challenges of adolescence, it is perhaps not surprising that adherence generally decreases as teenagers assume increasing responsibility for their medication-taking.

Other research on patient factors revealed findings that were often inconsistent. Some studies indicated that pediatric antiretroviral adherence was not related to gender, age, race, child’s knowledge of HIV status, structural social support, satisfaction with social support, or health status/virologic or immunologic factors. Other studies indicated that patient factors associated with better adherence were nonwhite (versus white) race, both younger and older age of child, child’s unawareness of his or her HIV diagnosis, beliefs regarding the positive impact of the medications on quality of life, lower intensity of alcohol use, housing stability, less depressive symptomatology and child stress, decreased child responsibility for medications, and improved health status/virologic or immunologic factors.

Family/caregiver factors are crucial to pediatric adherence, because infants and younger children depend almost entirely on a caregiver to administer medications. Their adherence to treatment, therefore, is largely determined by the resources and efficacy of their caregivers, who often struggle to administer medications to children who resist, refuse, or spit them out. Caregivers who are biological parents of HIV-positive children often share their diagnosis and confront challenges associated with their own illness and its comorbidities. Thus, they may be physically fatigued or debilitated. In these cases, treatment can become a reminder of the parents’ guilt about their role in their child’s acquisition of infection, which is yet another challenge to adherence.

Research on family/caregiver variables has indicated that children are more adherent if they receive their medications from foster parents rather than biological parents or other relatives. Parents of adherent children report higher perceptions of their ability to administer the prescribed doses and of the medication efficacy and less concern about others discovering their child’s diagnosis. The accuracy of caregivers’ information about the child’s HIV medications and corresponding dosages and dosing frequencies has been related to adherence, as has the number of barriers. In a study of 75 perinatally infected 3- to 13-year-olds in New York, Mellins et al found that nonadherence was related to worse parent-child communication, higher caregiver stress, lower caregiver quality of life, and worse caregiver cognitive functioning. However, in a sample of 30 infants and children under 12 years of age, Steele et al noted no association between adherence and parental perceived vulnerability or perceived barriers.

Overall, the literature on factors related to pediatric antiretroviral medication adherence (summarized in Table 2) has been largely atheoretical. Hammami et al, however, presented findings that confirmed the information-motivation-behavioral skills model. Specifically, in their qualitative study of 11 caregivers of HIV-positive children in Belgium, adherence was associated with (1) knowledge of the disease and its treatment; (2) motivation or willingness to adhere (which depended on the acceptance of the HIV diagnosis, the emotional quality of the mother-child relationship, and the recognition of future prospects); and (3) the capacity or ability to adhere (which depended on having the necessary cog-

### TABLE 2 Research on Factors Related to Pediatric Antiretroviral Adherence (Based on Quantitative Research With Tests of Statistical Significance)

<table>
<thead>
<tr>
<th>Significantly Associated With Adherence</th>
<th>Nonsignificantly Associated With Adherence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medication related</td>
<td></td>
</tr>
<tr>
<td>Twice-per-day (vs 3-times-per-day) nelfinavir regimen; shorter length of time since treatment initiation; nelfinavir rather than indinavir</td>
<td>Pill burden; drug toxicity</td>
</tr>
<tr>
<td>Patient related</td>
<td></td>
</tr>
<tr>
<td>Nonwhite (vs white) race; both younger and older age of child; children’s unawareness of their HIV diagnosis; beliefs regarding the positive impact of the medications on quality of life; lower intensity of alcohol use; housing stability; less depressive symptomatology; less child stress; decreased child responsibility for medications; improved health status/virologic or immunologic factors</td>
<td>Gender, age; race; child’s knowledge of HIV status; structural social support; satisfaction with social support; health status (virologic or immunologic factors)</td>
</tr>
<tr>
<td>Caregiver/family related</td>
<td></td>
</tr>
<tr>
<td>Foster (vs biological parent); higher self-efficacy; belief in the efficacy of the medication; less concern about hiding child’s diagnosis; better parent-child communication; less caregiver stress; higher quality of life; better caregiver cognitive functioning; better caregiver knowledge of antiretroviral medications; fewer barriers</td>
<td>Parental perceived vulnerability; perceived barriers</td>
</tr>
</tbody>
</table>
nitive and technical skills to follow a medication scheme, perceived self-efficacy, and problem-solving capacity).

**Evaluations of Strategies to Improve HAART Adherence**

Many studies of factors related to adherence also suggest strategies for improving HAART adherence, but few of these studies have been subject to empirical examination. We found 7 studies that focused on the evaluation of interventions to enhance antiretroviral medication adherence among pediatric populations, all of which were conducted in the United States. As described below, only 1 involved a randomized, control trial; the other reports presented primarily descriptive data on small samples.

Gigliotti et al. studied a DOT intervention, primarily to determine if prolonged elevation of VL could be attributable to poor adherence. They identified 6 perinatally infected children aged 3.3 to 11.5 years with elevated VLs for periods of months to years and for whom suboptimal adherence was suspected. DOT was administered in the hospital for 4 children and at an HIV program-sponsored summer camp for the other 2 children. DOT of 4 to 8 days lowered the VLs of all 6 of these heavily drug-experienced children with advanced HIV, exceeding a 90% decrease (1.0 log₁₀) in 4 of them. Surprisingly, as few as 4 days of DOT resulted in a 70% (0.5 log₁₀) drop in VL.

Roberts et al. examined an intensive DOT program among 6 families of children who showed continued high VLs for several years despite documentation of viral sensitivity to the prescribed medication regimen and caregiver assertions of regular medication administration. Their stepwise approach involved an initial referral to a home health nurse and then DOT supplemented by intensive education and counseling provided for patients and caregivers during a 4-day hospitalization and at 2 weeks postdischarge, and, if necessary, a physician-initiated medical-neglect report to state authorities. Although 5 of the 6 families cooperated with the visiting nurses, the referrals failed to result in sustained improvements in adherence. This lack of success was attributed to the many logistic problems of in-home visits and the lack of specialized training of the nurses. Four of the 6 patients responded to DOT, with sustained decreases in VLs. Medical-neglect reports did not improve adherence in the 2 other children, who demonstrated improvements in VLs and immune functioning after subsequent placement in foster care.

Lyon et al. had 23 HIV-positive black youths aged 15 to 23 years and 23 family members or other “treatment buddies” participate in a 12-week educational program. Six meetings included family members, and 6 meetings were with the youths only; all meetings were 2 hours and included a meal. Devices such as watches, pill boxes, and calendars were introduced. Three months postintervention, 91% self-reported improved adherence. Care managers corroborated these reports. None of the youths had significant improvements in VL (defined a priori as a 1-log reduction in VL or undetectable VL); however, 4 participants demonstrated improved immune functioning.

Via a retrospective medical chart review, Shingadia et al. located 17 perinatally infected children who had gastrostomy tube (GT) insertion for improvement of medication adherence at the median age of 2.9 years (range: 1.25–11.8 years). In the year after the procedure, all 17 patients were noted to be “adherent” by care providers, although a clinically significant improvement (ie, ≥2-log reduction in VL) was found in only 10 cases. The authors attributed this discrepancy to the initiation of a new medication regimen at the time of the procedure for all 10 of these patients, who showed improvement over the next year. They recommended, therefore, that the HAART regimen be changed after GT placement to minimize the impact of viral resistance secondary to nonadherence. Despite the intrusiveness of the medical procedure to insert the tubes, they were reportedly well tolerated by the children. Caregivers noted satisfaction with the devices, reductions in medication administration time, and improvement in child behavior during medication administration.

Rogers et al. evaluated the only theory-driven intervention, which was based on Prochaska's stages-of-change model. It involved an intensive 8-week program with 13- to 22-year-olds using videotape and audiotape material appropriate to assess stage of readiness to adhere with HAART: precontemplation, contemplation, preparation, action/maintenance, and relapse. The material was designed primarily for treatment-naive subjects to prepare them for successful initiation of HAART and long-term maintenance of adherence. Youths were to watch the videos with the study coordinators and then were offered the tapes to take home and share with family members. During the preparation program, participants practiced a regimen with surrogate pills for 1 to 2 weeks. Evaluation was hampered by high attrition (only 18 of the 65 enrollees completed the full program), which the authors attributed to difficulties in scheduling program visits and the labor-intensive nature of the intervention. Among these 18 participants, 78% moved forward in terms of their stage of readiness, and two thirds accepted HAART, with half of them maintaining adherence “most of all the time.”

Ellis et al. evaluated a clinical program that used home-based multisystemic family therapy delivered by mental health specialists, averaging 46 therapy sessions over 7 months. Children and families were referred to the program on the basis of their adherence problems or high VLs in the absence of resistance. Ninety percent accepted the referrals, and 95% received a full dose of treatment, which suggested high program feasibility. Among the 19 children aged 1 to 16 years who partici-
pated, caregiver-reported adherence did not change (likely because of ceiling effects according to the authors), but chart-abstracted VLs were found to significantly decrease from referral to the end of treatment. The mean change reflected a $>1$-log decrease in VL, and this change persisted through the 3-month follow-up as well.

In the only randomized, controlled trial, Berrien et al\(^5\) evaluated home nursing visits as a means of increasing adherence with 67 families. The home visits were designed to identify and resolve barriers to medication adherence, but specific barriers addressed by the intervention were not reported, with the exception of pill-swallowing training. In the treatment group, knowledge scores significantly improved, and self-reported adherence marginally improved.

In addition to these 7 published reports of study evaluations, we found 2 ongoing trials of adherence interventions, both of which were sponsored by the Centers for Disease Control and Prevention.

Pediatric Impact is a large multisite randomized, controlled trial of an intervention to improve adherence to antiretroviral therapy. It involves children aged 5 to 13 years and their primary caregivers in New York City and Washington, DC.\(^6\) The intervention includes a dedicated adherence coordinator, an initial needs assessment, and tailored modular interventions, including home-based services. Enrolled children and their caregivers are randomly assigned to either a “minimal” or “enhanced” arm and offered an individualized combination of the following 6 modules: HIV education, HIV diagnosis-disclosure education to children, behavior modification, medication swallowing, medication management, and referrals to social and mental health services.

Adolescent Impact, a randomized, controlled trial of an intervention to improve adherence to treatment and reduce sexual transmission risk behavior in youth aged 13 to 21, is currently ongoing in New York City, Baltimore, Maryland, and Washington, DC, at 5 clinical sites.\(^7\) The study includes adolescents with perinatally acquired infection and those who acquired their infection through sexual risk behavior. The intervention provides education, social support, and skills training through an integrated series of one-on-one and group sessions. Seven group sessions address health and developmental issues common to youth with HIV; 5 one-on-one sessions are used to tailor prevention messages to the unique health and risk profile of each teen. To aid adherence, participants receive organizational tools including a personal digital assistant with adherence software; an optional home visit is also provided.

**DISCUSSION**

In this qualitative systematic review of the literature on medication adherence in pediatric HIV, we located over 50 relevant studies, although many were purely descriptive and the empirical studies often included small samples. Researchers used a range of strategies for measuring adherence, with patient self-report or caregiver reports being the most common. In contrast to previous reports of suboptimal adherence, we found that estimates of adherence varied widely. Heterogeneous samples and methods complicated comparisons of adherence estimates both within and across studies. Correlates of adherence were reported atheoretically, generally as variables related to the medication, patient, or caregiver. Overarching models for understanding correlates and predictors of adherence were noticeably lacking, as were investigations of the role of societal and health care system factors. The literature on the development and formal evaluation of theory-driven strategies for improving adherence remains sparse. Overall, this review of the available literature leaves unanswered many key questions regarding adherence to antiretroviral medication in pediatric HIV infection.

**Recommendations for Assessing Adherence**

The literature across adult and childhood illness suggests each adherence-assessment method has advantages and disadvantages, with the trade-off generally assumed to be financial and logistic cost versus psychometric and epidemiologic accuracy.\(^8\) Reports from the patient, caregiver, or provider are the least expensive and most widely used methods, but they tend to overestimate adherence compared with other methods. Caregivers and patients are prone to social desirability and recall bias, and providers often misjudge adherence. However, 2 recent reviews demonstrated the validity of self-reports from adult patients, at least in terms of their associations with VL.\(^9\) Although EDM is often considered the gold standard and can provide useful data on timing and patterns of missed doses, it is subject to nonuse, and its high cost makes it impractical in most settings.\(^10\) Pharmacy refill data are clouded by the practice of some pharmacies to automatically mail medications at prescribed periods. Clinical variables such as VL are affected by numerous other factors such as the innate variability of disease activity and presence of resistant strains of virus.

Assessing adherence in pediatric populations poses specific challenges.\(^11\) The direct methods such as measurement of blood levels of medications are expensive, and the lack of information on the pharmacokinetics and pharmacodynamics of pediatric formulations of HAART medications renders interpretation of such data difficult.\(^12\) Even when relevant data are available, individual variations in metabolism and developmental changes complicate the task. In addition, the presence of measurable levels of a drug only means that the child took the last dose of that medication. Many pediatric formulations of HAART medications are available only in powder/liquid form, which restricts the utility of EDM.
and pill counts (although remaining powder and liquid medication can be quantified). Putting the liquid/power bottle within an EDM bottle has not been shown to be feasible. Prepackaged pill boxes or syringes further limit the use of EDM. Self-reports among children are subject to their cognitive abilities and developmental level. Caregiver report is subject to the vantage point of the individual making the report. Most children are cared for by many different caregivers. The parent or guardian, other family member, baby-sitter, home health aide, and school nurse each have different yet probably incomplete information about the specifics of medication-taking. No one individual is consistently the most reliable. For example, younger children may completely depend on a caregiver and know nothing about their prescribed regimens, whereas caregivers of older children play a much more peripheral role, perhaps just occasionally reminding the child to take a dose or administer a small percentage of doses.

The data from studies of pediatric HAART adherence fail to provide definitive guidelines or to identify any gold standard in terms of assessment methods. However, the limitations of any single assessment strategy suggest possible benefits of using multiple methods (or triangulation). Although methods other than self-report may be too unreliable (provider estimates), burdensome (telephone diaries), or costly (EDM) to use routinely in clinical settings, they might be considered for short-term use or to illuminate issues with respect to clinical care. Clinicians should remember that VL, which is often used as the criterion with which to compare adherence-assessment methods, is not a perfect indicator of adherence, especially in heavily and sequentially treated perinatally HIV-infected adolescents, in whom resistance occurs frequently.10

Face-to-face interviews can be administered quickly and easily and offer maximum flexibility in terms of asking about different classes of medications and times and amount of each dose.10 Simple inquiries about whether the child took the morning dose of the medication and then tracing backward to the evening and the day before can be followed by how many doses could not be given in the previous week. This or a simple question about how many doses were missed in the previous week may be an appropriate starting point.10,27,66,71 Questions regarding difficulties in administering medications such as child resistance or lack of disclosure of diagnosis to a temporary caregiver are useful as well. Providers should also elicit information regarding baby-sitting arrangements and methods used by the caregivers to remind others of the need for medication-taking.

A comprehensive assessment of adherence cannot be accomplished in 1 meeting but ideally is part of a structured program of adherence monitoring, allowing clinicians to identify nonadherence early and to focus on the families who have the most difficulty. Wiener et al10 advised assessing adherence during every visit and determining if nonadherence is confined to a particular drug or exists across medications. Reports of nonadherence provide opportunities to address problems and barriers. Dolezel et al11 suggested interviewing both the child and the caregiver, perhaps followed by a joint interview to reconcile any discrepant reports. Sensitivity to patients’ culture is key, according to DeRouen and Jackson,72 who recommended that providers examine their own personal biases; assess the patient’s background and cultural identification, including views of life, spirituality, death and dying, and the value of outside support systems; and be accepting, nonjudgmental, and flexible. Questions should facilitate reports of nonadherence (eg, “I understand it is very difficult to administer these medications to children: have you been able to give any to your child this week?”) rather than incriminating (eg, “We have talked about how important good adherence is. You have not missed any doses lately have you?”). Statements regarding how difficult it is to give all doses of medications in a timely fashion and maintenance of a nonjudgmental attitude throughout the interview are crucial. Assessing reasons for nonadherence (eg, work schedule, caretaking responsibilities for other family members, attempts to avoid thinking about the illness, caregiver’s own ill health) and patterns of nonadherence (eg, always missing the morning dose, missing doses on the weekends) is key to determining the appropriate intervention. It is important to remember that adherence to HAART is a dynamic and lifelong process. Crises, a new job, caregiver illness, a new relationship, a different baby-sitter, loss of housing, and a multitude of other factors that affect families can wreak havoc on medication administration in even the most adherent family.

Recommendations for Improving Adherence

Our review of the empirical literature on the efficacy of HAART-adherence interventions for pediatric populations located only 7 published studies, of which only 1 was a randomized, controlled trial. These studies provided some support for the utility and efficacy of DOT, a 12-week educational program, GT placement, and nursing home visits. However, most findings were based on pilot studies with very small samples, and adherence to the intervention itself was often problematic. The findings suggest that more intensive interventions are required to produce efficacious outcomes and that 1-time interventions in the absence of ongoing education and support may be insufficient. Overall, they exemplify the barriers that are often encountered in this challenging work. Although clinicians commonly assume that certain medications or regimens are easier to tolerate than others or rely on drug holidays to reinvigorate adherence efforts, the literature is sparse and inconsistent.
regarding the treatment-related factors that affect adherence. Many of these interventions have become clinical lore and require additional scrutiny.

With a few exceptions, much of the literature on strategies for enhancing adherence relies on clinical-practice wisdom.43,68 On the basis of work with adolescents and young adults, Dodds et al73 cited the benefits of adopting a developmental framework, fostering buddy systems, and determining psychological and psychosocial treatment readiness. They stressed the need for intensive, continuous, coordinated, and nurturing case management services, with screening and treatment for mental health problems and substance abuse. Program components that are important for adherence, in their experience, include colocation of key services, home visits, transportation, child care at home or in the facility, incentives, and reminder telephone calls and letters. In a review of the literature, Pontali et al74 noted the following patient and family strategies: setting up an adherence program, intensive education in preparation for HAART initiation (perhaps even hospitalization), use of reminders such as diaries and alarms, disclosure, inclusion of older children in the decision-making process, and DOT and GT insertion if necessary.

Given the lack of definitive work on medication-adherence strategies in pediatric HIV infection, clinicians might seek guidance from the research in other pediatric chronic illnesses, which share many of the challenges of HIV infection.63,64 Studies in the area of pediatric tuberculosis suggest that DOT may be a viable strategy, although there are important differences between treatments for the 2 diseases.75 Among children with asthma, juvenile rheumatoid arthritis, and type 1 diabetes, myriad adherence-intervention strategies have been examined.76 Educational strategies focus on providing written and verbal information to patients and their families about the illness and the regimen requirements, as well as the importance of adherence and anticipatory coping strategies for potential adverse effects. Organizational strategies emphasize simplifying regimens and changing clinic characteristics such as improving supervision of providers, decreasing waiting times, and boosting accessibility. Behavioral strategies incorporate visual reminders such as medication calendars and self-monitoring into the child’s daily schedule. They comprise encouraging parental support, dispensing incentives for desirable behaviors, and offering clinical interventions for the treatment of underlying behavioral and emotional difficulties in the patient and family. A review of the literature in these diseases revealed that “probably efficacious” strategies include organizational interventions, multicomponent packages, and operant learning and that behavioral, educational, and cognitive-behavioral strategies are “promising.”76

Recommendations for Future Research
Future research should focus on evaluating the validity and reliability of self-report adherence measures, because they are the most widely used and most practical tools for assessing adherence in clinic settings. Comparison between self-report and more objective measures and clinical indicators are needed to identify the most cost-effective assessment strategies.10 In addition, more data on the pharmacokinetics and pharmacodynamics of pediatric formulations of HAART medications will make therapeutic drug monitoring a more reliable method of assessment.

To investigate adherence levels and potential correlates of adherence, future research should incorporate prospective studies, perhaps beginning at HAART initiation and continuing throughout the disease course both in domestic and international settings. Steele and Grauer20 advised systematically examining factors that may predict adherence and developing risk profiles to target potential nonadherers. We need additional examination of factors related to caregivers (such as cognitive variables) and psychosocial adjustment and stress.16,38 Work is needed on the development and evaluation of theoretical models that incorporate multiple domains of influence, especially contextual factors. Models that have been applied successfully to HAART adherence among adult samples include those that are based on theories of stress and coping,77 social support,78 self-efficacy,79 and social problem-solving80; these may be worth investigating among pediatric samples. Qualitative data may facilitate the application of theoretical models to the complex context of pediatric HIV infection. For example, interventions will need to consider the normative developmental tasks faced by children with HIV infection as they survive into adolescence, including separation from adult figures and achieving a sense of mastery or control, which may incur risk-taking behavior that jeopardizes self-care with chronic illness. Theories that emphasize self-efficacy and social problem-solving are likely to be particularly relevant to adolescent developmental concerns and may be especially engaging if presented through high-technology modes such as the Internet.81,82

This research agenda will require multisite studies and the incorporation of non–US-based sites to increase sample size. The most efficacious interventions will likely need to go beyond merely cognitive intervention and education and counseling and will be long-term.45,68 Standardization of assessment methods and definitions of adherence in future work will facilitate the comparison of research results.

Critically important is the need to expand work in international settings and resource-constrained environments, where the pediatric HIV epidemic is expanding exponentially and the need for improved strategies for addressing adherence is most dire. Many developing countries have no health care infrastructure and no
models for the management of chronic diseases. The increased access to antiretroviral therapy in these areas needs to be coupled with strategies to address adherence. Research in the area of biomedical strategies may be especially relevant for clinical care in areas in which personnel to assist with adherence are scarce. Although empirical data are limited for many of these options, future work should investigate the effect of more palatable formulations, possibly liquids that do not require refrigeration or small-sized pills that need to be taken less frequently. More tolerable formulations, with fewer adverse effects, toxicities, and dietary requirements, also might be helpful.\textsuperscript{41} There is urgent need to develop innovative medication-delivery systems such as skin patches for antiretroviral therapy. These biomedical interventions may need to be combined with patient-focused educational and behavioral strategies and then reinforced by attention to larger structural issues such as access to care and decreasing stigma to effectively confront the challenges of pediatric HIV treatment.

**CONCLUSIONS**

Nonadherence to antiretroviral therapy has catastrophic consequences. Much work remains to be done to develop a meaningful multisystemic and effective yet cost-effective approach to assess and improve adherence to antiretroviral therapy for children with HIV infection.

**ACKNOWLEDGMENT**

This work was supported in part by Centers for Disease Control and Prevention cooperative agreement U64/CCU219450 (to A. Wiznia, principal investigator).

**REFERENCES**

1. Haynes RB. Compliance in Health Care. Baltimore, MD: Johns Hopkins University; 1979


43. Pontali E. Facilitating adherence to highly active antiretroviral therapy in children with HIV infection: what are the issues and what can be done? Paediatr Drugs. 2005;7:137–149


57. Ellis DA, Naar-King S, Cunningham PB, Secord E. Use of multisystemic therapy to improve antiretroviral adherence and health outcomes in HIV-infected pediatric patients: evaluation of a pilot program. AIDS Patient Care STDs. 2006;20:112–121


63. Rapoff MA. *Adherence to Pediatric Medical Regimens*. New York, NY: Kluwer Academic/Plenum; 1999


Between Stillness and Story: Lessons of Children’s Illness Narratives

Sayantani DasGupta, MD, MPH
Division of General Pediatrics and Program in Narrative Medicine, Columbia University, New York, New York

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

ABSTRACT

In answer to an increasingly impersonal medical environment, educators in the medical humanities frequently turn to narrative studies to teach students for an emotionally fulfilling and interpersonally related professional practice. However, to elicit, to interpret, and to integrate patient stories into their work effectively, physicians must be in a state of awareness and attention, attuned to their emotional and intellectual reactions. The experiences of children and their families, in the form of pediatric illness narratives, hold unique insights for physicians in how to engage in an ethical, empathetic, and self-reflective practice. In particular, these narratives demonstrate the importance not only of story but also of stillness or silence to the practice of medicine. The voices of patients and their families hold both literal and allegorical lessons for physicians in how to move toward a medical practice involving not only diagnosis and treatment but also recognition and healing.
YEARS AGO, I attended a talk at Columbia University by the renowned Buddhist scholar Robert Thurman. It was in a packed lecture hall, with a few hundred individuals in attendance. Before beginning his talk, Thurman had us do something that I have never seen done before in an academic environment; he had us sit in silence.¹ We should be still, he told us, dutifully counting our breaths, closing our eyes, and being quiet. We should, in short, allow ourselves to simply be. The experience of sitting silently in a meditative stillness allowed the audience to leave aside external concerns and attend Thurman’s lecture in a full and present way. Although I have neither the renown, nor perhaps the courage, to ask such a thing of a live audience, I ask you, my readers, to sit with this story for a moment before you continue with this essay.

Breathe in.
Breathe out.

Gong. (Thurman simulated the sound, confessing that he had not been mindful enough to remember to bring a bell.)

As such stories are, this story has been told here for a reason. As Thurman used his lecture to suggest a central role for meditation in higher learning, this essay will similarly suggest a central role for stillness, what Thurman might call inner silence, in the work of caring for the ill, the work of medicine. Claude Debussy once said that music is the space between the notes.² To pose it a little differently, if the beauty of a sonata emerges from the balance of notes and silence, then this essay posits that the heart of medicine beats in the space between stillness and story.

The modern health care crisis has been framed in many ways as a crisis of story. Shortened visit times, increased reliance on technologic tools for diagnosis, and a lack of focus on the individual physician-patient relationship has created a medical environment fraught with dissatisfaction and frustration for both ill individuals and their providers. Medical educators have turned increasingly to the humanities, particularly narrative studies, to bring storytelling back to the center of health care. One prevailing perspective in medical humanities is that being able to understand the patient story in its fullness (ie, metaphor, frame, plot, and point of view) is the avenue through which to enter more fully into patient stories and thus a more ethical, empathetic, and satisfying professional practice.³ Narrative-based educators teach medical trainees close reading of classic literary texts (from Dostoyevsky and Chekhov to James Joyce and Henry James) that deal with the universal human experience but not necessarily the experience of suffering and illness. Other humanities educators have turned to autobiographical narratives authored by individuals suffering from illness and disability, as well as biographical narratives authored by their family members and caregivers. By learning to witness the textural voices of those affected directly by illness and disability, students learn a skill parallel to that they will use in their future health care practices. Students begin to ask themselves what it is to experience suffering, what it is to represent that experience (in written text or oral story), and what it is to be a witness (professional or familial) to the experience of suffering.

One aspect of this renewed medical conversation regarding story is an investigation into the act of listening itself. What is it to be a witness to suffering? What skills are required for physicians to elicit, to receive, and to interpret their patients’ illness stories effectively? In his classic work, A Fortunate Man: The Story of a Country Doctor, John Berger describes the work of the physician as that of recognition. “This individual and closely intimate recognition is required on both a physical and psychological level,” he writes. “On the former it constitutes the art of diagnosis. Good general diagnosticians are rare, not because most doctors lack medical knowledge, but because most are incapable of taking in all the possibly relevant facts—emotional, historical, environmental as well as physical. They are searching for specific conditions instead of the truth about a man which may then suggest various conditions.”⁴ This sort of listening demands a professional voice “shows the subject as actor” as opposed to the passive voice, which “denotes whether the subject performs or receives the action expressed by the verb.”⁵ The active voice “shows the subject as actor” as opposed to the passive voice, which “shows the subject as acted on.”⁶ Here, then, is one small beacon illuminating the crisis of story in medicine. If subjecthood in the medical profession is predicated on assuming an active voice in most professional activities (doing things), then it only follows that medical subjects, namely, physicians, should approach the witnessing of stories from a similar stance. We speak of “getting the story,” as if it were an object to be found and fetched intact, an active, even athletic, process of discovery, archaeology, and search and rescue. However, the witnessing of suffering is a process of being “acted on,” humbled, changed, and filled in addition to being informed. This sort of listening demands a radical shift in stance, in grammatical voice, such that physicians not only act but also allow themselves to be acted on.

Consider this popular quotation, which has been at-
tributed to the Buddha, “Don’t just do something. Stand there.”9 Whether coined by the Buddha or by a T-shirt maker, these words do provide a revised understanding of active and passive voice. “Standing there,” being still, takes on new meaning in this light, as a different deeper kind of doing, that is, an “un-doing.” an opening, a laying bare. Much more than merely “do no harm,” passivity in the medical relationship can, with this understanding, become an active process whereby stillness facilitates for the physician-listener a deeper understanding of the story being told. Consider the words of Ralph Waldo Emerson, who writes, “When I watch that flowing river, which, out of regions I see not, pours for a season its streams into me, I see that I am a pensioner; not a cause, but a surprised spectator of this ethereal water; that I desire and look up, and put myself in the attitude of reception, but from some alien energy the visions come.”10 The insight of Emerson’s visions emerge from looking to the “ethereal water” with an “attitude of reception,” as medical insight may emerge from a similar posture that allows physicians to wade in and become filled with the streams of their patients’ stories. In an essay titled “Intelect,” Emerson writes, as if about medical thinking, “Our thinking is a pious reception. Our truth of thought is therefore vitiates as much by too violent direction given by our will, as by too great neg- licence. We do not determine what we will think. We do not decide what to do. We just act, and see what happens.”11 This endeavor in passive voice is not in lieu of the active work of medicine but in addition to or perhaps in balance of it. In the words, again, of Emerson, “It seems as if the law of the intellect resembled that law of nature by which we now inspire, now expire the breath; by which the heart now draws in, then hurls out the blood, the law of undulation. So now you must labor with your brains, and now you must forbear your activity, and see what the great Soul shoveth.”11 We breathe in so that we can breathe out. 

The question is how to teach transcendental objectives such as witnessing, attentiveness, recognition, and “pious reception.” Given this problem of educating physicians in the skills of deeper listening, I imagine that Robert Thurman might suggest, in that unerringly contradictory way of Eastern mystics, that we teach students to listen to stories by first learning to listen to silence. 

Physicians are particularly accustomed to hearing and interpreting silent stories, such as the surgeon who interprets the story told by the anesthetized bleeding body, the rheumatologist who hears the inarticulate story told in symptoms of pain, or the geriatrician who daily witnesses the silent story of death. Pediatricians, I posit, are more accustomed than most to negotiating the fine balance between silence and story. Indeed, our work incorporates a daily practice of listening to at least outward silence. Charon writes:

We clinicians . . . act almost as ventriloquists to give voice to that which the patient emits. I put it this way because the patient cannot always tell, in logical or organized language, that which must be told. Instead, these messages come to us through the patient’s words, silences, gestures, facial expressions, and bodily postures as well as physical findings, diagnostic images, and laboratory measurements, and it is our task to cohere these different and sometimes contradictory sources of information so as to create at least provisional meaning.8

Pediatricians feel in our core this lesson that Charon helps us articulate. We know that much of the pediatric story emerges from nonverbal communication on the part of the child. Very young children are developmentally unable to access and to use language in the same manner as adults, but even older children are often unable to articulate their experiences, because of their social voicelessness. Despite efforts on the parts of pediatricians to elicit and to hear the voices of their patients, stronger still are the nonverbal messages transmitted to children through their parents’ expectations of docility and obedience as integral to good public behavior, adults’ obvious preference to speak directly to one another, and the very situation of being small, ill, and disrobed in a room with a stranger to whom one’s parent has inexplicably conferred the power to gaze on, to probe, and to invade one’s body. For adolescents, voicelessness may be a manifestation of internal emotional struggles. Consider the words of Jenn, a 15-year-old girl with anorexia. “I like to be alone,” says she, “There are a lot of expectations, and you kind of let those go for a while. When I’m alone, I feel like I can be sad. I don’t talk about my feelings to other people. I keep them to myself. I don’t know myself that well. I definitely care about what other people think about me; that’s a huge part of me, so that could be it. I like to be alone, then I don’t have to know, I just am.”12p71 From the prelinguistic damaged infant in the NICU to the frightened, distrustful, or sullen teen in the adolescent clinic, pediatricians become skilled at constructing stories from silence, as if creating matter from what was once ether.

Recognizing and interpreting the silence of others is not the same, however, as training students in mindful practice, attentiveness, or the ability to listen to one’s inner silence. Jenn, the adolescent of the earlier example, retreats inward, to a place where she can freely identify and experience herself through her emotions. It is this inner place of self-observation that we in narrative medicine are trying to help our students reach. It is not emotional detachment from one’s patient or a sentimental overattachment that makes a good physician but speech and action that emerge from an awareness of one’s inner landscape. “Speaking about it, you kind of have to know how you feel,” explains Jenn. Both speak-
ing and acting as a doctor require knowing how one feels. This emotional self-awareness cannot, of course, overwhelm the listener or transgress into emotional self-indulgence. Inward retreat cannot be an escapist venture, as Jenn’s experience seems to be, away from demanding human relationships, because a physician’s work is predicated on such interactions. In a surprisingly profound conclusion that sounds a bit more like Eastern philosophy than teenage angst, Jenn states that when she is alone she does not have to know, but just is. This is the difficult part. How do we hyperintellectual physicians get from that place of knowing to a place of just being?

During that talk long ago, Robert Thurman suggested that the way to construct an undergraduate education that created responsible global citizens was not simply to suggest but actively to teach students the skill of looking mindfully inward. Buddhism and Western education, Thurman posited, share the goals of investigative and penetrative knowledge, because Buddhism considers understanding and knowledge liberating forces. He quoted his friend and teacher, the Dalai Lama, as being concerned with the fact that Western education relies solely on developing the intellect in this process, whereas a Buddhist education also focuses on developing “the good heart.” The rest of Thurman’s talk took up this challenge by the Dalai Lama for Western education to focus on the “good heart” as well as a good intellect. Basing his comments on Buddhist teachings, whereby an ultimate presence in the body is seemingly contradictorily used to transcend the body, whereby an inward solitary gaze is used to recognize one’s connections to the universe, whereby the still passivity of the meditative stance enables the most active state of being, Thurman suggested that one answer to the crisis in education might be to turn vast university gymnasiums into communal meditation chambers, that educators should look beyond a classical humanities education to train pupils in humanity. In Buddhism, as in many Eastern religions, meditation becomes the means to channel one’s focus and to achieve a deep inner knowledge.1 Meditation breeds recognition of self, and self-recognition enables recognition of others. Although extending Thurman’s argument to medicine and turning our anatomy and histology laboratories into meditative rooms may not be a practical solution, we may imagine that one answer to the crisis of modern health care, the crisis of story, lies in teaching all medical students the skills at the heart of meditation and self-recognition, skills central not only to eliciting and witnessing our patient’s stories but also to integrating them fully into our own consciousness.

This is the point at which we medical humanities educators might throw up our hands and turn over our responsibilities to chaplaincy, religious studies, or the meditative disciplines. Although I encourage closer ties with such colleagues, this essay also suggests a role for narrative studies, and scholars of illness narratives, in helping students negotiate the relationship between stillness and story. Here is where my own argument takes a Thurman-esque contradictory turn. If Thurman might suggest that we teach students to listen to stories by first teaching them to listen to silence, then my own work with illness narratives suggests that the best way to teach a student to listen to silence might be through reading stories.

Illness narratives (autobiographical or biographical stories written by those suffering illness and their caregivers) represent a genre of writing that has grown significantly in the past few decades. Sociologist Arthur Frank has suggested that such stories, or pathographies, are a postmodern phenomenon, in which narratives authored by the ill give voice to an experience that was once narrated solely by the medical establishment. In other words, illness stories speak where there once was silence, giving voice to what once was inarticulate.13 Narratives written by the ill hold more than lessons about the experience of illness; they also hold lessons directly applicable to the work of inward recognition and outward witnessing. The reason for this is twofold, attributable to both the nature of stories themselves and the particular position of the authors of such stories. Regarding the first, it is again Emerson who can help us understand the power inherent in narration:

Each truth that a writer acquires is a lantern, which he turns full on what facts and thoughts lay already in his mind, and beheld, all the mats and rubbish which had littered his garret become precious. Every trivial fact in his private biography becomes an illustration of this new principle, revisits the day, and delights all men by its piquancy and new charm. Men say, Where did he get this? and think there was something divine in his life. But no; they have myriads of facts just as good, would they only get a lamp to ransack their attics withal. We are all wise. The difference between persons is not in wisdom but in art.11

In the process of narrating their stories, authors of autobiopathographies or biopathographies hold lanterns up to illuminate the experiences of their lives and give them an existence that both belongs to and transcends the idiosyncratic author. In addition, illness stories are inherently social and bring attention to the relationships of tellers to listeners, sufferers to caregivers, and patients to physicians. Frank writes, “Any person’s story depends on others who become less other as the enmeshment of stories teaches interdependence. I confess to believing that in learning this interdependence patients have a qualified advantage over clinicians. This advantage may have less to do with physical suffering . . . and more to do with not having a particular institutional face that must be sustained before one’s colleagues.”14 These stories are steeped, then, in lessons of relationship and
interdependence, lessons directly relevant to the physician’s work.

Pediatric pathographies, published stories about children’s illness experiences, are a particularly small subgenre of this writing, perhaps because of some of the same issues of social voicelessness discussed earlier, as well as children’s lack of access to formal written language and publishing. Although published narratives regarding children’s illness experiences tend to be written by either adults remembering their childhood illness experiences or the parents of ill children, they are no less the narratives of the children about whom they are written, in the sense that all illness stories are cowritten, some by caregiver and patient, some by sufferer and loved one, and some by present and past selves. These stories make clear not only that children transmit a nuanced, often silent, story that their physicians must mindfully attend but also that ill children and their families have much to teach us about the process of attention itself. The 2 particular narratives I visit in this discussion were chosen because they hold lessons that are both literal, giving insight into the experience of being an ill child, and allegorical, in that these stories seem to hold some lessons central to an ethical, empathetic, and spiritual (or perhaps centered) practice of medicine.

Let us consider first Lucy Grealy’s memoir of her childhood experience with Ewing’s sarcoma of the jaw, *Autobiography of a Face.* Grealy’s narrative, which was first published when she was in her early thirties, is simultaneously a reflection back to her childhood cancer and a rumination on the lifelong effects of that cancer and the subsequent, seemingly endless, operations to reconstruct her face. Although Grealy does locate the source of her identity in her face, writing, “my face, my self,” it would be a mistake to consider this text a mere reversal of the Narcissus myth, a girl’s quest to find her face in the mirror. She writes:

> “Even as people confirmed that this was now my face, even as people congratulated me, I felt I was being mistaken for someone else. The person in the mirror was an imposter—why couldn’t anyone else see this?”

In reading this, my breath is taken away, not only by the beauty of Grealy’s writing and her power to transport me into the body of an ill 9-year-old girl but also by the resonance of her images to the work of medicine itself. The ethical work of medicine is situated in the very sort of recognition with which Grealy struggles, the recognition of the face that is not the self, the face on the other side of the mirror, the face of the suffering patient. In his discussion of the work of philosopher Emmanuel Levinas, medical educator Craig Irvine writes, “Like Narcissus, medicine is stupid with, in thrall to its own reflection . . . [but] medicine is a moral endeavor, commanded by ethics, by transcendence, by an infinity beyond its comprehension.” What lies beyond the comprehension of the self, Levinas teaches us, is the other, and the “calling into question of the same” by the presence of the other is what Levinas calls ethics. Irvine here helps deepen our understanding of Berger’s “recognition,” Charon’s notion of “attention,” even Thurman’s Buddhist mindfulness. Ethics precedes even knowledge, according to Levinas, and the primordial ethical act is to answer the call of other’s suffering, to see what Levinas terms the other’s “face.” In Grealy’s text, it is her own changed image that calls into question the sameness of that which she calls self, and her quest to recognize her own face is the ethical struggle in which she engages.

Grealy’s narrative can be read as a manifestation of medicine’s ethical struggle, but she also illustrates the skills necessary for “recognition” to be enacted. Grealy’s youthful self exemplifies the attentive state central to physicians’ ethical work, even as her own physicians seem barely to see her. In the following passage about Grealy’s first visit to the oncologist, it is young Lucy, rather than her doctor, who is still, vessel-like, filled with the environment and people around her, taking in the minutest details of her world. She writes:

> “When we were finally in Dr Woolf’s office . . . we encountered his telephone, apparently a permanent appendage. He could carry on a conversation with my mother, me, his nurse, the secretary down the hall, and someone on the phone simultaneously, he had it down to an art. His manner was gruff and unempathetic . . . . His office was as drab as the waiting room but saved by a large, multipaned window that looked out onto a well-tended courtyard with banks of blue flowers and ivy-clad trees. I spent a lot of time forcing myself to look out that window, because even on that first visit I knew that room was no place for me. . . . The first examination . . . I was asked to strip down to my underwear, which I did, feeling humiliated and exposed. While the doctor talked to the nurse, my mother, and the person on the phone tucked beneath his chin, he prodded me with his hands, hit me just slightly too hard with his reflex hammer, and spoke far too loudly. When he touched me, I could feel the vibrations of his voice in my own chest, feel them lapping through my body’s cavity the same way you feel a car passing too closely.”

Here, medicine is practiced the wrong way around, with the child patient emptying herself and being filled with the reverberations of medicine’s boorish care. Repeatedly, Grealy’s narrative references the notion of silence and disappearance. After her first chemotherapy treatment, her father asks her, “That wasn’t so bad now,
was it?” The 9-year-old Grealy does not answer, and her adult counterpart writes, “Speaking seemed like something one would grow tired of.”15(p77) The ill child becomes the mirror that shows the world around her. It is her own otherness, brought on by medicine, that causes her to doubt her very existence as a part of this world she now reflects.

Like meditation, where stillness of the body encourages alertness of the mind and spirit, Grealy’s enforced external stillness or silence is accompanied by great internal attention. Her silent observations of her outside world are brought forth by her increasing self-presence in her inner world. Illness is the medium through which external stillness or silence is accompanied by great internal alertness of the mind and spirit, Grealy’s enforced stillness or silence is accompanied by great internal alertness of the mind and spirit, creating friction and space by rubbing against the viscera, the muscles of my stomach, my back, my lungs. Grealy describes being able literally to feel her internal anatomy lesson. I had never known it was possible to feel your organs, feel them the way you feel your tongue in your mouth, or your teeth. My stomach outlined itself for me; my intestines, my liver, parts of me I didn’t know the names of began eating up, trembling with their own warmth, creating friction and space by rubbing against the viscera, the muscles of my stomach, my back, my lungs.16(p73)

Grealy’s narrative has obvious literal lessons for those of us who deliver care to either children or adults. We must attend our patients with our entire beings, Grealy teaches us, and thereby make space for them, their fear, their voice, and their presence, in the rooms of medicine. If we read Grealy’s narrative on a more-allegorical level, however, gleaning lessons about our own doctoring, then we begin to discover one path to achieving our ethical goal of fully present care. The attention that the physician brings to the clinical encounter comes not from losing ourselves, abandoning our intellectual, emotional, or cultural senses of self and overidentifying with our patient, but from being fully present as listeners. As Grealy describes being able literally to feel her internal self, this sort of witnessing on the part of physicians comes from being fully aware of one’s emotional viscera, one’s affective anatomy. What stories are easier for us to hear because they remind us of, or do not remind us of, our personal and family histories? Which patients evoke unarticulated anger or frustration? What personal emotional needs affect our clinical care? Grealy’s narrative suggests that we physicians must receive the stories around us not through an emptying of the inner self but through a heightened self-awareness that does not go down the road of solipsism but helps facilitate an ultimate state of inner stillness. The ability to engage in such self-reflective, attentive practice is not just an ideal but is a necessity for the work of doctoring. Without it, as we see from Grealy’s narrative, we physicians risk transferring to our patients the burden of recognition, making them into the mirrors that reflect back to us our own narcissistic reflections.

Narratives of pediatric illness written by parents can similarly shed light on an ethical, attentive practice of medicine. In Asian religion scholar Sam Crane’s memoir about his son, Aidan’s Way,17 the relationships between stillness, spirituality, and interrelatedness are explored. Crane’s son Aidan is literally and socially voiceless, in that he is so damaged from anoxic birth trauma that he is unable to see, to walk, or to speak, he is confined to a wheelchair, and he must contend with a seizure disorder and multiple other difficulties. Crane’s narrative is simultaneously about a parent giving voice to the experiences of his voiceless child and about the profound lessons this child silently speaks to his parent and, through the parent’s narrative, to the physician-reader. Crane formulates his narrative, and his title, as a Taoist quest, which is a search for the Tao or “The Way.” In Crane’s words, “The Way is, simply put, the complex unity of nature. It is not a transcendent God standing above and apart from His creation; it is more like a common, earthbound origin from which all things grow and are sustained.”17(p47) Crane’s tale is interspersed with ancient Chinese fables and scriptures, all of which resonate with lessons of inward attention and universal interrelatedness.

Aidan’s very existence raises the questions that we are accustomed to thinking of as those of medical ethics. Crane is called on repeatedly to justify the meaning of Aidan’s life to family, friends, community, and the insurance companies that cover Aidan’s medicines and his expensive nutritional supplements. After one such confrontation, Crane is able to reflect on the broader social, ethical, and spiritual implications of these questions:

The whole affair . . . [was] representative of the broader domination of utilitarian thinking. Aidan was costly, his care was using up finite resources that might have gone to some other good purpose. For what the company spent on him, several other cases might be supported. The greatest good for the greatest number could arguably be better served by cutting off Aidan . . . each dollar had to be made to go farther, money could not be wasted on just one profoundly mentally retarded boy who would never walk or talk or see anyway. It’s their fault, the parents, for keeping these kinds of kids alive.

It’s not just money that distorts our view of human worth. Social status, cultural attainment, physical beauty: all of these and more creep into our calculations of an individual’s value. These sorts of criteria are so common-place that it sometimes seems remarkable when we are reminded that none of them fully capture the possibilities of personhood. And that is what Aidan does. His value comes precisely from the challenge he poses to the usual definitions of “value.” He is a living reminder that
the range of human experience is broader than the narrow confines of balance sheets and business plans. Without a word, he poses the deepest questions. What is life? What makes any life, even one so limited, worth it? Strangers have come up to us on crowded streets, touching his shoulder or tousling his hair, giving us their abbreviated answers. Usually they say something about love or grace, something well beyond the material concerns of everyday life. We are constantly reminded of these more sublime things because, with Aidan, it's never about utility or efficiency or productivity, it's about humanity.17(p247)

Crane uses the Taoist metaphor of the useless tree to describe Aidan's life. He writes, “The central image is a gigantic tree, gnarled and knotty, with rotting wood and fetid leaves. It is apparently worthless, devoid of alluring fruit or durable timber. Attracting little attention, it has grown unencumbered, spreading out its branches so that it could shelter a thousand teams of horses in its shade. It stands in silent denial of our obsessions with the useful, the productive, the efficient, the worthy.”17(p269) We see in one of the interpretations of the useless tree, dare I say, useful correlation to the work of medicine. In one version of the story, the enormous tree stands next to, and in its stillness shelters, a village shrine, “becoming, through its impressive immensity, a part of the shrine itself.”17(p269) So too can the attentive listener, the mindful physician, stand next to the shrine that is the patient’s story and, through his or her very being, his or her ability to resonate and to reverberate from the chants emanating from the shrine, become part of it.

Like the physician who functions as the still vessel, Aidan’s stillness indeed invokes the stories of those around him. After the initial questions, Aidan’s mainstream classmates accept him completely into their community. Crane writes, “Their abilities were magnified in the mirror of his limitations, so they were pleased to congregate around him.”17(p149) Crane tells the story of a little boy with a speech impediment in Aidan’s class who is too shy to speak in front of the other able-bodied children but befriends Aidan, chatting away to him because Aidan is silent, still, and nonjudgmental, the best sort of listener. Aidan’s relationships with Crane’s self but changes it with his very presence, pulling Crane into his own silence, just as he changes his classmates’ perceptions of their own abilities, “magnifying” them through his own limitations. Crane’s narrative, which is Aidan’s narrative as well, helps clarify not only why stillness works but also how it works. Stillness, or passive voice, in medicine, becomes active in its very relational abilities, in its ability not only to stand next to but also to become part of the stories of suffering.

It is thus that the stories of children like Lucy and Aidan intertwine with physicians’ own stories, get into our bones, and provide us with an opportunity to learn something deeper about our work and ourselves. There is a wisdom emanating from these children’s illness stories, and this wisdom is perhaps the key to an emotionally centered and spiritually fulfilling practice of medicine. Arthur Frank has suggested that we should not only think about illness stories, analyzing and critiquing their literary qualities, but also think with them, gleanings truths about our own lives in relation to these narratives.13 In thinking with pediatric illness narratives such as those of Lucy and Aidan, we physicians have the potential to tap into a knowledge that precedes our technical and academic skills. Many Eastern religions do not consider adulthood as the highest stage of development. Rather, “the end of the cycle is that of the independent, clear-minded, all-seeing Child. That is the level known as wisdom. And so Eastern texts such as the Tao te Ching urge us, ‘Return to the beginning; become a child again.’”2(p151) Fully entering narratives of children’s illness experiences, thinking with such stories, is one way to return to this place.

In a story called, “The Surgeon as Priest,” Richard Selzer introduces a healer skilled in this sort of wisdom. Selzer writes about Yeshi Dhonden, the personal physician to the Dalai Lama, a saffron-robed, shorn, “golden” man who “receives” his patient through his presence and his touch, the palpation of her pulse. Through his wordless presence and his laying on of hands, the Buddhist healer not only diagnoses his patient but also engages in an act of deep spiritual recognition. “So!” writes
Selzer, “Here then is the doctor listening to the sounds of the body to which the rest of us are deaf. He is more than a doctor. He is a priest.”

During the process of his diagnosis, both Yeshi Dhonden and his patient are wordless; their storytelling and story-listening occur through silence. Although Selzer’s narrator does not abandon his Western medical training, he suggests that his future work in healing will be infused with the spiritual lessons of the monk healer. “Now and then it happens, when I make my own rounds, that I hear the sounds of his voice, like an ancient Buddhist prayer, its meaning long since forgotten, only the music remaining. Then a jubilation possesses me, and I feel myself touched by something divine.”

Physicians have the rare privilege of not only caring for patients but also allowing them to become our teachers, schooling us in attention, awareness, presence, interdependence, and empathy. It is an act of profound humility to turn the tables of power and to learn at the feet of those one would teach, to listen to the voices and silences of the children and adults for whom we care, and, in doing so, to approach a more-mindful medicine. This sort of reciprocal practice, listening to and in doing so learning to listen to our patients, is an act that propels physicians one step further on the path of ethical work. There reverberates in the background a deeper lesson, one that it would be foolhardy to try to name. Instead, let me turn, as Crane does, to the words of Chang Tzu (the bold is mine):

> Joy and anger, sorrow and delight, hope and regret, doubt and ardor, diffidence and abandon, candor and reserve: it’s all music rising out of emptiness, mushrooms appearing out of the mist. Day and night come and go, but who knows where it all begins? It is! It just is! If you understand this day in and day out, you inhabit the very source of it all.

When I am with a patient, I try to remember sitting in that room, in that moment before Thurman ended our silence, and I try to listen very carefully, to the story before me and to the stillness behind me. Sometimes, if I am very lucky, I hear them all, the monk, the healer, the child, and the philosopher, lifting their voices, music out of emptiness, mushrooms out of mist, singing. “We are! We just are!”

**ACKNOWLEDGMENTS**

I thank Rita Charon, Maura Spiegel, and Craig Irvine for their invaluable editorial contributions to this article. In addition, I thank Dr Jayne Rivas and the pediatric department at St Vincent’s Catholic Medical Center Manhattan, whose request for a grand rounds presentation on pediatrics and spirituality generated a much earlier version of this essay.

**REFERENCES**

17. Irvine CA. The other side of silence: Levinas, medicine, and literature. *Lit Med.* 2005;24:8–18
The Role of Mometasone Furoate Aqueous Nasal Spray in the Treatment of Adenoidal Hypertrophy in the Pediatric Age Group: Preliminary Results of a Prospective, Randomized Study

Marco Berlucchi, MD, Daria Salsi, MD, Luisa Valetti, MD, Giovanni Parrinello, MD, Piero Nicolai, MD

*Department of Pediatric Otorhinolaryngology, Spedali Civili, Brescia, Italy; †Department of Otorhinolaryngology and ‡Section of Medical Statistics, University of Brescia, Brescia, Italy

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

ABSTRACT

OBJECTIVE. We evaluated the efficacy of mometasone furoate aqueous nasal spray in decreasing adenoid size and reducing the severity of chronic nasal obstruction symptoms in children affected by adenoidal hypertrophy.

METHODS. Sixty children were recruited in a 2-stage, randomized, placebo-controlled trial. All patients complained of chronic nasal obstruction symptoms, and nasal endoscopy showed >75% choanal obstruction attributable to adenoid pads. In the first stage, 30 patients (group A) underwent mometasone treatment (50 μg per nostril per day) for 40 days, and 30 children (group B) received placebo. In the second stage, at the end of the first 40-day treatment period, patients in group A who showed subjective and objective clinical improvement were divided into 2 subgroups; group A1 (11 children) received topical intranasal steroid treatment on alternate days for the first 2 weeks per month, whereas group A2 (10 children) continued daily mometasone treatment for the first 2 weeks per month. After 3 months, all children were reassessed.

RESULTS. Fifty-seven children completed the study according to the protocol. After the first treatment period, the severity of symptoms and adenoid size decreased for 21 patients (77.7%) in group A. No improvement was observed in the placebo group. After 3 months of additional therapy, group A2 patients demonstrated a more-pronounced reduction in adenoid size compared with group A1 patients. No statistically significant change in symptoms was identified. Mometasone treatment was well tolerated by all patients.

CONCLUSIONS. Mometasone furoate aqueous nasal spray may be considered useful in decreasing adenoid pad size and the severity of symptoms related to adenoidal hypertrophy. Children with adenoidal hypertrophy that is not associated with tonsillar hypertrophy should be considered for intranasal mometasone treatment before surgery is planned.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-1769
doi:10.1542/peds.2006-1769

Key Words
adenoidal hypertrophy, adenoid pad, topical steroids, nasal steroids, mometasone furoate

Abbreviations
AH—adenoidal hypertrophy
OSAS—obstructive sleep apnea syndrome
MF—mometasone furoate

Accepted for publication Nov 28, 2006

Address correspondence to Piero Nicolai, MD, Department of Otorhinolaryngology, University of Brescia, Piazza Spedali Civili 1, 25123 Brescia, Italy. E-mail: pieronicolai@virgilio.it

PEDiATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
**Methods**

**Study Subjects**

Between February and July 2004, 60 children affected by AH and referred for exclusive adenoidectomy were recruited at the Department of Pediatric Otorhinolaryngology, Spedali Civili of Brescia (Brescia, Italy). At enrollment, patients needed to meet the following inclusion criteria: (1) adenoid pad occluding ≥75% of the nasopharynx, as determined with nasal endoscopy; (2) age between 3 and 7 years; (3) symptoms consistent with AH lasting ≥12 months; and (4) no previous adenoidectomy. Children with comitant tonsillar hypertrophy, positive history of allergy or atopy, upper respiratory infection within the past 2 weeks, nasal anatomic anomalies (eg, nasal septum deviation) or sinonasal diseases such as hypertrophy of inferior turbinates and/or nasal polyposis, craniofacial malformations including clefts, genetic diseases (eg, Down syndrome), neurologic disorders, cardiovascular diseases, immunodeficiency, history of epistaxis, hypersensitivity to steroids, or intranasal, topical, or systemic steroid or antibiotic treatment within the past 4 weeks were excluded.

**Study Design**

The design of the study and the informed consent form for this 2-stage, prospective, placebo-controlled, randomized study were approved by the institutional review board of the Spedali Civili of Brescia (Brescia, Italy). Before initiation of treatment, informed consent for participation in the study was obtained from the parents or legal guardian of each child recruited. In the first stage, after a primary clinical evaluation and nasal endoscopy (time 0), children were assigned randomly to receive a single intranasal administration in each nostril of either MF aqueous nasal spray (50 μg) (group A) or a placebo saline solution nasal spray (group B) for 40 days.

At the end of the first 40-day period (time 1), both groups were reassessed, to evaluate the efficacy of treatment. Patients in group A who showed improvements in clinical findings and decreases in adenoid pad size, such that adenoidectomy could be avoided, were considered responders. In the second stage, responders underwent “maintenance therapy” for 3 months and were divided randomly into 2 subgroups; patients in group A1 received intranasal MF treatment on alternate days for the first 2 weeks per month, whereas those in group A2 continued daily topical steroid treatment for the first 2 weeks per month. Nonresponders and placebo-treated patients, for whom neither subjective nor objective improvement was apparent, were scheduled for adenoidectomy. After 3 months (time 2), the 2 subgroups were reassessed.

**Assessment and Patient Treatment**

At baseline, each child underwent physical evaluation and nasal endoscopy; clinical history was obtained from the parents with a questionnaire. Patient history included age, gender, weight, history and family history of atopy or allergy, and use of drugs. Symptoms such as nasal obstruction, rhinorrhea, cough, snoring, and obstructive sleep apnea were also evaluated at recruitment and at each subsequent visit, by using a clinical scoring system ranging from 0 to 3 (0 = absent; 1 = occasional; 2 = frequent; 3 = daytime and nighttime symptoms). In this way, we assigned to each aforementioned symptom a score related to severity. All scores were summed to obtain an overall symptom score for each patient. During follow-up visits, parents reported on their children’s symptoms and eventual adverse effects (eg, nasal bleeding).

Nasal endoscopy was performed to estimate adenoid pad size and to identify other sinonasal disorders. After local anesthesia administration and decongestion of the nasal mucosa for 10 minutes with cotton pledgets soaked in phenylephrine and mepivacaine hydrochloride, endoscopic examination was conducted by using a rigid (2.7-mm diameter) or flexible endoscope, according to the compliance of the child. All nasal endoscopies, which were performed when the patient was performing quiet nasal breathing, were recorded by using a Karl Storz camera (Karl Storz, Tuttingen, Germany). Pictures of both choanae were transferred to a computer. On each picture, the degree of adenoid obstruction was estimated as a percentage of the posterior choanal area occupied by...
adenoid tissue. Children underwent endoscopic evaluation at time 0, time 1, and time 2.

Compliance with drug administration was assessed biweekly in telephone interviews with parents. Moreover, the occurrence of any disease and related therapy were recorded; the use of systemic steroid therapy resulted in exclusion of the patient from the study.

Statistical Methods
Baseline characteristics are presented as median and interquartile range or percentile, according to the type of variables analyzed. Possible baseline imbalance between the 2 groups was analyzed by means of the Wilcoxon exact test or Fisher’s exact test. The primary end point of the first stage (reduction of AH) and the effects of treatment on secondary end points (score variations for chronic nasal obstruction symptoms) were analyzed with the Kruskal test or Wilcoxon exact test. Prognostic factors for responsiveness were studied with logistic regression. Potential differences at the end of the second stage were evaluated with the Kruskal test or Wilcoxon exact test. A Bonferroni correction for multiple comparisons was also applied. A statistical value of \( P < .05 \) was considered significant in all comparisons.

RESULTS
Sixty children were enrolled in the study and during the first stage were assigned randomly to receive MF (group A; \( n = 30 \), 18 male subjects and 12 female subjects; median age: 5 years) or placebo (group B; \( n = 30 \), 13 male subjects and 17 female subjects; median age: 4 years). No patient had a personal or family history of allergy or atopy, had undergone previous surgery, had received any drugs in the past 4 weeks, or had immunodeficiencies.

All children in group B and 27 patients in group A completed the study. Among the 3 patients in the latter group who did not complete the study, 1 was withdrawn for having received systemic steroid therapy for acute ethmoiditis, 1 was lost to follow-up monitoring, and 1 dropped out of the study.

At time 0, there were no significant differences between study groups with regard to demographic features or symptoms such as rhinorrhea, cough, and snoring, whereas nasal obstruction and obstructive sleep apnea were significantly more severe in group A. The mean overall symptom scores were 11 for group A (range: 6–15) and 10 for group B (range: 4–15). Nasal endoscopy was well tolerated by all children. No complication was observed during the initial diagnostic evaluation or after nasal endoscopies. Neither anatomic sinonasal abnormalities nor inflammatory changes of nasal mucosa except for AH were found with rhinoscopy. At recruitment, the mean choanal obstruction was 88.5% in group A (range: 75%–100%) (Fig 1) and 76.5% in group B (range: 75%–100%) (Table 1).

After the first treatment period (time 1), 21 (77.7%) of 27 patients who received MF were classified as responders, whereas 6 (22.3%) of 27 were nonresponders. Symptoms attributable to chronic nasal obstruction improved significantly in group A (mean overall symptom score: 3), whereas they were substantially unchanged in group B (mean overall symptom score: 9). The adenoid size also decreased significantly (Fig 2) in patients given MF (mean choanal obstruction: 64%) (Table 2). Only 1 patient in the steroid group reported episodic epistaxis, which resolved spontaneously without treatment. No other adverse effects were reported. At time 0, the only statistically significant difference between responders and nonresponders was related to obstructive sleep apnea (Table 3).

In the second stage, 11 of 21 responders were assigned randomly to group A1 and 10 to group A2. All children completed this stage of the study and, after 3 months of maintenance therapy, no significant differences in symptom scores were observed between the 2 subgroups, whereas the mean choanal obstruction in group A2 was less than that in group A1 (56% vs 65.5%; \( P < .001 \)). Compliance with therapy during the last 3-month period was satisfactory in both subarms, and no adverse effects were observed.

DISCUSSION
AH is probably the most frequent pathologic condition occurring in the pediatric age group. It leads to different clinical manifestations according to adenoid size. Bilat-
eral nasal obstruction is a primary complaint that can be associated with different sleep disorders, ranging from snoring to OSAS.5 In such a situation, which is often observed when palatine tonsillar hypertrophy is also present, patients typically complain of both nighttime and daytime behavioral illnesses (ie, intermittent sleep, sleepwalking, morning headaches, difficulty concentrating, sleepiness, enuresis, slow feeding, and poor growth), which may lead to cardiorespiratory syndromes such as cor pulmonale in extreme cases. Rhinorrhea, mouth-breathing, hyponasal speech, and cough can also be observed in patients with AH. Moreover, nasal blockage and auricular secretions attributable to the adenoid pad relationship with nasal fossae and eustachian tubes may be present and represent a source of infection. In addition, AH seems to favor the occurrence of recurrent and effusive otitis media and recurrent and chronic rhinosinusitis.

Finger palpation, transoral mirror adenoid examination, and baseline lateral soft-tissue radiographs of the nasopharynx commonly have been used to assess adenoid size. However, these techniques are now considered inaccurate, and nasal endoscopy has been established as the method of choice for assessment of AH.6–9 In recent decades, technologic advances have led to the development of small-diameter (2.7-mm), flexible and rigid endoscopes that allow accurate nasal endoscopic examinations with no complications. As observed in the present study, rhinoscopy performed after positioning of cotton nasal pledgets soaked in topical anesthetic can be considered a thorough, safe, well-tolerated, reproducible procedure for assessment of adenoid pads. Furthermore, it provides information about the nasal mucosa, nasopharyngeal secretions, and the presence of other sinonasal anatomic abnormalities or diseases.

At present, AH is one of the most frequent indications for surgery in childhood, and adenoidectomy commonly is considered definitive treatment for nasopharyngeal obstruction. Nevertheless, this surgical technique has been the subject of some criticism. Paulussen et al10 hypothesized that the removal of adenoid lymphatic tissue could have a negative impact on the systemic immunologic system. Moreover, immediate postoperative or late bleeding is observed in ~1% of children who undergo adenoidectomy. Furthermore, it is well demonstrated that adenoids may recur after surgery in 10% to 20% of cases.11,12

In the past decade, several authors have proposed the use of topical nasal steroids to decrease AH, with the intent to preserve immunologically active tissue and to avoid the anesthesiologic and surgical risks inherent in adenoidectomy.1–4 In 1995, Demain and Goetz1 described an adenoid reduction using an aqueous beclomethasone solution for 17 children with AH. All patients complained of classic symptoms of chronic nasal obstruction and had an estimated ≥90% adenoid obstruction of the choanae, as determined with primary nasal endoscopy. In that study, patients were divided randomly into 2 groups; the first group received 4 weeks of intranasal aqueous beclomethasone (336 μg/day) nasal spray treatment, followed by 4 weeks of placebo

| TABLE 1 Baseline Demographic, Symptom, and Choanal Obstruction Data |
|-----------------------------|-----------------------------|-----------------------------|
|                             | Group A (N = 27)            | Group B (N = 30)            | P    |
| Age, median (interquartile range), y | 5.0 (3.5–5.0)             | 4.0 (3.0–5.0)              | .388 |
| Male, n                     | 16                          | 13                          | .23  |
| Weight, median (interquartile range), kg | 18 (16–20)                | 19 (15–20)                | .975 |
| Nasal obstruction score, median (interquartile range) | 3.0 (2.5–3.0) | 2.0 (2.0–3.0) | .0202 |
| Rhinorrhea score, median (interquartile range) | 2 (1–3)                  | 2 (1–2)                    | .468 |
| Obstructive sleep apnea score, median (interquartile range) | 1 (0–2)                   | 0 (0–1)                    | .0113 |
| Cough score, median (interquartile range) | 1.00 (1.00–3.00)          | 1.00 (0.25–2.00)          | .0973 |
| Snoring score, median (interquartile range) | 3.0 (2.0–3.0)            | 2.0 (1.3–3.0)             | .350 |
| Choanal obstruction, median (interquartile range), % | 90 (85–90)               | 80 (75–90)                | .13  |
treatment, whereas the second group underwent the same treatments in the reverse order. After the first 8-week period, all children continued topical nasal administration of beclomethasone (168 μg/day) for 4 months. At the end of the study, significant improvements in adenoid obstruction and symptoms were observed for all patients. Of the 17 subjects, however, 7 had a clinical history of atopy (asthma, allergic rhinitis, or atopic dermatitis) and 12 had a family history of atopy. In our opinion, the characteristics of those patients could have influenced the final outcomes.

In 2001, Brouillette et al2 tested the efficacy of another intranasal steroid treatment for OSAS in a randomized, triple-blind, placebo-controlled, parallel-group trial investigating the use of fluticasone propionate nasal spray versus placebo for 25 children affected by OSAS, as demonstrated with polysomnography. Thirteen of 25 patients underwent topical intranasal fluticasone therapy (50 μg of active drug) with 1 spray per nostril twice daily for the first 7 days and then once daily for an additional 5 weeks. The remaining 12 children received placebo. After treatment, the mixed/obstructive apnea/hypopnea index, frequency of hemoglobin desaturation, and respiratory movement/arousals decreased more in the fluticasone-treated group, compared with the placebo-treated group. Moreover, improvements in symptom scores were observed for 69% of children who received fluticasone.

In 2003, Criscuoli et al,3 studying 53 children, reinforced the conclusions reached by Demain and Goetz.1 For the first time, they reported on the long-term outcomes of treatment with aqueous nasal beclomethasone for patients with adenotonsillar hypertrophy. Twenty-four patients exhibited improvement after 2 weeks of steroid treatment, and an additional 24 weeks of therapy at a lower steroid dose maintained clinical improvement at 52 and 100 weeks for 45.8% of those patients.

Recently, Cengel and Akyol4 assessed the efficacy of MF in the treatment of AH, in a prospective, controlled, randomized, clinical trial. Of 122 patients enrolled, 67 received intranasal MF (100 μg/day) therapy for 6 weeks, whereas 55 patients were assigned to the control group. After treatment, a significant decrease of the adenoid mass was observed for 67.2% of the study group, whereas the clinical situation was unchanged in the control group. It is noteworthy that 8.9% of MF-treated patients had a positive history and prick-test results for atopy.

Among several commercially available steroid nasal sprays (beclomethasone dipropionate, budesonide, flunisolide, fluticasone propionate, MF, and triamcinolone acetonide), we chose to test MF for 4 reasons, namely, (1) the drug had been reported previously not to cause any adverse tissue changes in the nasal mucosa of patients treated for long periods,13 (2) it has no effects on growth in children,14 (3) it has no effects on the hypothalamic-pituitary-adrenal axis,15 and (4) the systemic availability of the drug after topical administration is lower than that of other steroids.16 To date, no standard indications regarding dosage and duration of topical intranasal steroid therapy for the treatment of AH have been established. Compared with the aforementioned trials, we chose to administer a lower daily steroid dose in each nostril but for a longer time (first treatment

### TABLE 2
Chronic Nasal Symptoms and Choanal Obstruction After the First 40 Days of Treatment

<table>
<thead>
<tr>
<th></th>
<th>Median (Interquartile Range)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A (N = 27)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal obstruction score</td>
<td>1.000 (0.000–2.000)</td>
<td></td>
</tr>
<tr>
<td>Rhinorrhea score</td>
<td>1.00 (0.00–1.50)</td>
<td>.041</td>
</tr>
<tr>
<td>Obstructive sleep apnea score</td>
<td>0.000 (0.000–1.000)</td>
<td>.0034</td>
</tr>
<tr>
<td>Cough score</td>
<td>1.00 (0.00–2.50)</td>
<td></td>
</tr>
<tr>
<td>Snoring score</td>
<td>2.00 (0.00–2.00)</td>
<td></td>
</tr>
<tr>
<td>Choanal obstruction, %</td>
<td>20.0 (12.5–32.5)</td>
<td></td>
</tr>
<tr>
<td>Group B (N = 30)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal obstruction score</td>
<td>0.000 (0.000–0.000)</td>
<td>.00005</td>
</tr>
<tr>
<td>Rhinorrhea score</td>
<td>0.00 (0.00–1.00)</td>
<td>.041</td>
</tr>
<tr>
<td>Obstructive sleep apnea score</td>
<td>0.000 (0.000–0.000)</td>
<td>.0034</td>
</tr>
<tr>
<td>Cough score</td>
<td>0.00 (0.00–1.00)</td>
<td></td>
</tr>
<tr>
<td>Snoring score</td>
<td>0.00 (0.00–0.00)</td>
<td></td>
</tr>
<tr>
<td>Choanal obstruction, %</td>
<td>0.0 (0.0–0.0)</td>
<td>.007</td>
</tr>
</tbody>
</table>

### TABLE 3
Demographic, Symptom, and Choanal Obstruction Data for Nonresponders Versus Responders at Time 0

<table>
<thead>
<tr>
<th></th>
<th>Nonresponders (N = 6)</th>
<th>Responders (N = 21)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, median (interquartile range), y</td>
<td>4.0 (3.3–4.8)</td>
<td>5.0 (4.0–5.0)</td>
<td>.282</td>
</tr>
<tr>
<td>Male, n</td>
<td>3</td>
<td>13</td>
<td>.601</td>
</tr>
<tr>
<td>Weight, median (interquartile range), kg</td>
<td>17 (14–20)</td>
<td>18 (16–20)</td>
<td>.491</td>
</tr>
<tr>
<td>Nasal obstruction score, median (interquartile range)</td>
<td>3 (3–3)</td>
<td>3 (2–3)</td>
<td>.108</td>
</tr>
<tr>
<td>Rhinorrhea score, median (interquartile range)</td>
<td>3.0 (1.5–3.0)</td>
<td>2.0 (1.0–2.0)</td>
<td>.103</td>
</tr>
<tr>
<td>Obstructive sleep apnea score, median (interquartile range)</td>
<td>2 (2–2)</td>
<td>0 (0–2)</td>
<td>.0361</td>
</tr>
<tr>
<td>Cough score, median (interquartile range)</td>
<td>2.5 (1.3–3.0)</td>
<td>1.0 (0.0–3.0)</td>
<td>.217</td>
</tr>
<tr>
<td>Snoring score, median (interquartile range)</td>
<td>3 (3–3)</td>
<td>3 (2–3)</td>
<td>.288</td>
</tr>
<tr>
<td>Choanal obstruction, median (interquartile range), %</td>
<td>90 (86–90)</td>
<td>90 (85–95)</td>
<td>.633</td>
</tr>
</tbody>
</table>
period: 40 days). A single low dose of intranasally administered steroid was well accepted by parents/legal guardians and children, thereby increasing compliance. Moreover, only 1 complication (episodic epistaxis) was observed, which demonstrates the safety of intranasal MF administration. Although group A had more severe clinical findings than group B at time 0, 77.7% of children treated with MF showed significant improvements in symptoms and significant reductions in adenoid size after the first 40 days of treatment, thus avoiding adenoidectomy. Interestingly, among the analyzed factors (ie, age, gender, weight, symptoms, and choanal obstruction), only obstructive sleep apnea was statistically significant in discriminating between responders and nonresponders. However, it is difficult to attribute the true value to these results because of the small number of patients studied. Although at our knowledge there are no publications in the literature supporting or criticizing a low topical steroid dosage for the treatment of AH, we modified the treatment schedule after the first treatment period, with the intent to maintain the outcomes while decreasing the dosage of the intranasal MF aqueous nasal spray further. Because we could not modify the steroid dose delivered in each puff, we decided to modify the duration of treatment, thus increasing the compliance of the children and their families. For this reason, responders were divided into 2 subgroups and underwent 2 different maintenance therapies. At the end of the second stage of treatment, group A2 showed significant reduction of AH, compared with group A1, whereas symptoms were comparable. Therefore, daily administration of MF for 2 weeks per month, in addition to maintaining successful clinical results, seems to decrease AH further. It is critical to highlight that we obtained these successful results for children with only AH. By obstructing the postnasal space, adenoids prevent steroids from acting on the palatine tonsils. In the study by Demain and Goetz, no evident tonsillar changes were observed for 7 children who showed, besides AH, moderate tonsillar hypertrophy. Because the steroid nasal spray acts especially in the nasal fossa and nasopharynx, we tested the effects of intranasal MF therapy on patients affected exclusively by AH, with no tonsillar hypertrophy. Several mechanisms, such as direct lympholytic action, inhibition of inflammation, and alteration of adenoid bacterial flora, have been suggested to explain how steroids decrease adenoid pad volume and improve symptoms of AH, although none has yet achieved widespread acceptance.1,17

CONCLUSIONS
We report a prospective, randomized, placebo-controlled trial on the efficacy of MF aqueous nasal spray versus placebo for the treatment of AH in children. Topical intranasal MF therapy can be considered a good therapeutic option to decrease AH. Nasal administration of this steroid is safe, reproducible, easily performed, and well tolerated by pediatric patients. Daily use for 2 weeks per month after an initial 40-day-period treatment seems to be the ideal maintenance schedule. Obviously, the indications for adenoidectomy remain unchanged for nonresponders. Future studies addressing the long-term (>1-year) efficacy of MF therapy and the identification of factors that could be used to select nonresponders are warranted.

REFERENCES
COMMENTARY

Go Ahead Punk, Make My Day: It’s Time for Pediatricians to Take Action Against Media Violence

Victor C. Strasburger, MD

Department of Pediatrics, University of New Mexico School of Medicine, Albuquerque, New Mexico

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

News flash! Researchers have just found that element "X" contributes between 10% and 30% to the occurrence of heart disease. What happens next? Well, of course, the manufacturers of element X claim that there is no link and that the research is spotty anyway; Hollywood writers and producers deny that element X ever appears in any of their films, and if it does, it is never "gratuitous," nor can anyone interfere with their right to show it, which is guaranteed in the First Amendment; and pediatricians continue to counsel parents about whether car seats should face forward or backward but apparently could not care less about element X.

How would have thought that 50 years after the first congressional hearings on media violence that we would still be debating whether it contributes to real-life violence? As one leading group of researchers has stated, there should no longer be any controversy that a cause-and-effect relationship exists.1 More than 3500 reports, including 1000 research studies, have been made; <30 studies have found no relation.2 In fact, the connection between media violence and real-life aggression is nearly as strong as the connection between smoking and lung cancer.3 Given the complexities of performing social-science research, this finding is quite striking.4 Not everyone who smokes will get lung cancer, and not everyone who views media violence will become aggressive. In addition, as the authors of a study published in this month’s Pediatrics Electronic Pages4 and the National Television Violence Study5 pointed out, context is extremely important, as are mediating factors such as personality differences and parenting. Overall, an estimated 10% to 30% of violence in society can be attributed to the impact of media violence.6 Clearly, media violence is not the leading cause of violence in society, but it is a significant one, and one that could be altered more easily than other root causes.7

Hollywood needs to understand that the science is convincing and that an intelligent response is needed. No one is suggesting censorship. However, violence in movies has steadily been increasing during the past decade, and not only have the ratings not kept pace, but they have underrated the increasing violence—so-called ratings creep.8 Movies that were R-rated a decade ago are now PG-13.9 Webb et al8 found remarkable levels of violence in recent PG-13 films (87% overall), with 7 films containing >100 acts of violence and firearms used in nearly one third of all films. Most often, the violence is portrayed as being “justified,” which is the single strongest positive reinforcer known.5,7 The notion of justifiable violence leads nations into wars, in case anyone was wondering about the significance of all of this. So Hollywood writers and producers have to be much more careful about how they portray violence, in what context, and to whom. The Motion Picture Association of America needs to open its board to the light of public scrutiny (its members’ identities are a closely guarded secret) and add a pediatrician and a child psychologist or communications expert.

Parents need ratings that they can understand and use, and a universal ratings system for all media would go far toward helping them.4 However, parents also have to stop thinking that sex is “worse” than violence in movies. It is not.10 The United States is the only country
Parents also need to be cautious when allowing their young children to view PG-13 movies; when I go to R-rated movies, invariably there are a few 7-year-olds sitting in the audience. We need a scientific study of why parents take young children to movies that are inappropriate for them, and pediatricians need to counsel parents that such films may be frightening and harmful.11

Finally, pediatricians need to be more sensitive to the effects of all media on children and adolescents. While pediatricians are spending their precious few minutes of office counseling on car seats and bicycle helmets, the media may represent a far greater threat to the health of young people. Virtually every concern that pediatricians and parents have about children has some root in the media (eg, violence, sex, drugs, academic performance, obesity, suicide, eating disorders).12 Pediatricians can no longer afford to dodge the bullet of media influence!

REFERENCES
10. Strasburger VC. Risky business: what primary care practitioners need to know about the influence of the media on adolescents. *Prim Care.* 2006;33:317–348
Acute Myopericarditis After Multiple Vaccinations in an Adolescent: Case Report and Review of the Literature

Maria T. Thanjan, MD\textsuperscript{a}, Prema Ramaswamy, MD\textsuperscript{b}, Wyman W. Lai, MD, MPH\textsuperscript{a}, Irene D. Lytrivi, MD\textsuperscript{a,b}

\textsuperscript{a}Department of Pediatric Cardiology, Mount Sinai Medical Center, New York, New York; \textsuperscript{b}Division of Pediatric Cardiology, Maimonides Medical Center, New York, New York

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

ABSTRACT

We report a case of postvaccination acute myopericarditis in an adolescent. The patient presented with acute chest pain, diffuse ST-segment elevation, and elevated cardiac enzyme levels. Cardiac MRI was consistent with acute myocarditis. He recovered within a few days with nonsteroidal antiinflammatory treatment and remains clinically stable, with improvement of MRI findings at the 10-week follow-up. Postvaccination cases of myopericarditis reported in the pediatric literature are also reviewed.

CARDIAC COMPLICATIONS AFTER routine immunizations are extremely rare, especially in children. The vaccination that has received great attention recently is that for smallpox, particularly after reinstallation of the vaccination for military personnel in 2002 and the reports of >50 cases of probable myocarditis temporally related to it.\textsuperscript{1} A prospective Finnish study identified electrocardiogram (ECG) changes suggestive of myocarditis without other evidence of cardiac disease in 3% of military recruits after vaccination against mumps, polio, tetanus, smallpox, diphtheria, and type A meningococcus.\textsuperscript{2} Few cases of myocarditis after smallpox vaccination have been reported in the pediatric literature as well.\textsuperscript{3–8} Other vaccinations have also been implicated in children, albeit much more rarely. There is 1 reported case of myocarditis that developed hours after diphtheria-tetanus-acellular pertussis (DTaP) vaccination in a 3-month-old\textsuperscript{9} and another case of myocarditis after tetanus vaccination alone in a 14-year-old.\textsuperscript{10} Our case is the first report of myopericarditis temporally related to booster DTaP administered simultaneously with hepatitis A and meningococcal vaccines before college entry in an adolescent.

CASE REPORT

A 17-year-old white boy was admitted to the PICU with intermittent, retrosternal chest pain for 1 day. He had no history of angina pectoris, cocaine abuse, hyperlipidemia, or congenital heart disease. He also had no history of autoimmune disorder, no previous travel, and no previous flu-like illness with diarrhea and/or respiratory symptoms suggestive of an infectious syndrome. Two days before admission, he received DTaP, meningococcal conjugate (MCV4), and hepatitis A vaccines. This was the first time he had ever received meningococcal or hepatitis A vaccines. The following day he complained of diffuse arthralgia and chest pain, and he had a low-grade fever. In the emergency department his physical examination was normal, but he was found to have diffuse ST-segment elevation on ECG (Fig 1A) and elevated serum cardiac enzyme levels: troponin I, 15.9 ng/mL (reference: <0.4 ng/mL); creatine kinase, 1106 U/L (reference range: 24–145 U/L); and creatine kinase MB, 44.4 U/L (reference range: 5.7–16.6 U/L). A chest radiograph was normal, and a two-dimensional echocardiogram showed no segmental wall-motion abnormalities, normal biventricular systolic function, and no pericardial effusion. Other laboratory findings included a white blood cell count of 8.3 × 10\textsuperscript{9}/L (with 62% neutrophils, 22% lymphocytes, and 3.7% eosinophils), hemoglobin

Key Words: immunization, myopericarditis, adolescence

Abbreviations: ECG, electrocardiogram; DTaP, diphtheria-tetanus-acellular pertussis; Ig, immunoglobulin

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2605
doi:10.1542/peds.2006-2605

Accepted for publication Nov 28, 2006
Address correspondence to Irene D. Lytrivi, One Gustave L. Levy Place, Box 1201, New York, NY 10029. E-mail: renalytrivi@yahoo.com
PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
level of 14.6 g/dL, C-reactive protein level of 59 mg/L (reference: <8 mg/L), erythrocyte sedimentation rate of 17 mm/hour (reference: <15 mm/hour), C4 level of 38 mg/dL (reference: 16–47 mg/dL), and C3 level of 142 mg/dL (reference: 88–201 mg/dL). Rectal and nasopharyngeal swabs were sent for adenovirus and enterovirus cultures, and the results were negative. Lyme immunoglobulin G (IgG)/IgM antibody, cytomegalovirus IgG, Epstein-Barr virus capsid IgM antibody, antinuclear antibody, and anti-DS DNA screens were also negative. Convalescent samples for the above-named infectious etiologies were not obtained.

Over the following 24 hours, the patient’s cardiac enzyme levels continued to rise to a maximum of 22.8 ng/mL for the troponin I. Gadolinium-enhanced cardiac MRI was positive for myocarditis, as evidenced by delayed subepicardial enhancement of the basal to midanterolateral and inferolateral regions of the left ventricular wall (Fig 2 A and B). Chest pain was relieved with ketorolac, and his serum enzyme levels gradually decreased. The patient was discharged 4 days later on indomethacin 50 mg orally 3 times per day with the recommendation of restricting physical activity for a total of 4 to 6 weeks. Serum cardiac enzyme and C-reactive protein levels had all normalized 1 week after the onset of the symptoms. ECG at that time showed normalization of the ST segment (Fig 1B), and at the 3-week follow-up, there was diffuse T-wave inversion (Fig 1C). At 10 weeks, ECG findings had resolved, and the follow-up cardiac MRI showed decrease in the area of contrast enhancement (Fig 2 C and D). The patient remained asymptomatic with normal ventricular function as assessed by echocardiogram and MRI.

**DISCUSSION**

Although most cases of myocarditis in the United States and Western Europe result from viral infections, other etiologies such as other infectious agents, various drugs, hypersensitivity, and autoimmune disorders have been occasionally implicated. In our case, the possibility of a viral etiology cannot be definitely excluded; however, the negative viral serology and the absence of symptoms make this a less likely possibility. The proposed mechanism of myocardial injury in cases reported after vaccination is a hypersensitivity reaction.10,11 Hypersensitivity myocarditis is usually a retrospective circumstantial diagnosis that is suspected because of the temporal link between receiving the vaccination or other offending agent and onset of symptoms. Pathogenesis is related to a maladaptive immune response that leads to myocardial injury, as evidenced by biopsy specimens in cases of myocarditis after smallpox vaccination that have re-

---

**FIGURE 1**
A, Twelve-lead ECG at presentation showing ST elevation and PR depression on I, aVL, and V1–V6. B, One week later there was ST normalization. C, At the 3-week follow-up, there was T-wave inversion on I, aVL, and V1–V6.

**FIGURE 2**
A and B, Cardiac MRI short-axis and 4-chamber views of the left ventricle in the acute phase showing the areas-of-contrast delayed enhancement (arrows). C and D, Equivalent views at follow-up showing a decrease in the area of involvement.
vealed CD3+ T-cell infiltrate with prominent degranulating eosinophils.12

Patients with myopericarditis can be asymptomatic or present with chest pain, shortness of breath, palpitations, fatigue, and decreased exercise tolerance. Troponin I elevation has high specificity in supporting the diagnosis, especially in conjunction with strong clinical suspicion.13 On the other hand, erythrocyte sedimentation rate and C-reactive protein are more sensitive than specific as general markers of inflammation and are probably useful in monitoring the benefit of antiinflammatory treatment.

ECG abnormalities accompanying cardiac enzyme level elevation in the appropriate clinical setting are highly suggestive of myopericarditis and are reported in up to 90% of cases.14 Of interest is the natural progression of ECG changes that is commonly associated with myopericarditis and was observed in our patient. Four stages of evolution have been described.15 The initial stage occurs during the first few days and lasts up to 2 weeks. The most characteristic ECG finding is ST-segment elevation and PR-segment depression. This phase is followed by ST-segment normalization and T-wave flattening; this stage is extremely variable in duration, lasting anywhere from a few days to several weeks. Stage III usually occurs in the second or third week of illness and can persist for a few weeks. During this time, full T-wave inversion occurs. In the final stage, there is resolution of the T-wave abnormality and a return to baseline normal ECG findings. It is important to note that although the 4-stage progression of the ECG findings is “pathognomonic” for myopericarditis and very useful for differentiating it from other clinical syndromes such as acute myocardial infarction or benign early repolarization, in many instances only part of the sequence occurs, and atypical presentations are common.

With cases like ours, in which suspicion of myopericarditis is high and risk for coronary artery disease is low, a noninvasive test to rule out ischemic heart disease is an acceptable alternative to cardiac catheterization. Cardiac MRI has been shown to be highly sensitive and specific for myocarditis detection and differentiation from myocardial infarction.16,17 Delayed gadolinium enhancement of the myocardium is usually subepicardial (as opposed to the subendocardial involvement usually seen with ischemia and infarction) and can be focal or diffuse depending on the timing after the onset of symptoms.16,18 The cellular-level mechanism proposed to explain this contrast enhancement is rupture of the myocyte membrane and diffusion of the contrast agent into the necrotic cells, which results in increased tissue-level concentration.19

Because most reported cases of postvaccination myopericarditis are self-limited, the role of treatment is not clear. On the basis of limited data and mostly anecdotal evidence, patients are either not treated or nonsteroidal antiinflammatory drugs12,20 and corticosteroids6,21 are used in cases of more severe clinical presentation (evidence of congestive heart failure). No data exist regarding long-term sequelae, and as a consequence, recommendations for long-term follow-up of these patients are not evidence-based. The role of myocardial damage, shown by contrast cardiac MRI, as a predictor for long-term outcome needs to be investigated further.

CONCLUSIONS
Our case highlights the fact that pediatricians should be aware of the often-dramatic presentation of postvaccination myopericarditis and its usually benign clinical course. The diagnosis of myocarditis should be entertained when acute-onset chest pain is accompanied by ECG changes and elevated cardiac enzyme levels. In cases in which the above-described presentation is temporally related to routine immunizations, the immunizations should be considered as a possible underlying etiology.

REFERENCES
Serum Osmolal Gap in Patients With Idiopathic Nephrotic Syndrome and Severe Edema

Gaurav Kapur, MD, Rudolph P. Valentini, MD, Abubakr A. Imam, MD, Amrish Jain, MD, Tej K. Mattoo, MD, FRCP

Division of Nephrology and Hypertension, Children’s Hospital of Michigan, Wayne State University School of Medicine, Detroit, Michigan

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

ABSTRACT

Pseudohyponatremia in idiopathic nephrotic syndrome with severe edema is attributed to hyperlipidemia that results in displacement of a portion of water phase of plasma. Current methods of measurement of serum electrolytes are unaffected by hyperlipidemia. In this report we demonstrate that patients with idiopathic nephrotic syndrome with severe edema and true hyponatremia may have an increased rather than normal osmolal gap. We believe that this could be secondary to non-Na\(^+\) and non-K\(^+\) osmoles in response to plasma-volume contraction secondary to hypoalbuminemia. This observation has implications for management of severe edema in such patients, because fluid restriction could increase their risk for pre–renal failure.

Hyponatremia is a common observation in patients with idiopathic nephrotic syndrome (INS) with severe edema. In view of coexisting hyperlipidemia in such patients, hyponatremia has traditionally been labeled as pseudohyponatremia rather than true hyponatremia. Pseudohyponatremia is the method-dependent artifactual reduction of serum sodium attributed to measurement error caused by hyperlipidemia, which results in displacement of a portion of plasma water by increased lipid concentration. Currently, direct potentiometry is the method of choice for measuring blood electrolytes. With this method, undiluted serum samples are used to measure transmembrane potentials resulting from electrolyte gradients by using an ion-selective electrode. Thus, the water content of the sample does not affect measurement of serum sodium, and pseudohyponatremia is rarely seen.\(^1\) True hyponatremia in patients with INS is included in the category of diseases with extracellular fluid volume expansion, as in congestive heart failure (CHF), and cirrhosis. In CHF and cirrhosis there is plasma-volume expansion but effective circulatory volume depletion secondary to primary reduction in cardiac output or vascular resistance. However in INS, hypoalbuminemia leads to plasma-volume contraction. Edema is secondary to compensatory retention of sodium and water by the kidney because of effective circulatory volume depletion in CHF/cirrhosis and plasma-volume contraction in INS.\(^2\) Studies have reported increased serum osmolality and osmolal gap (defined as the difference between the measured and the calculated serum osmolality) attributable to non-Na\(^+\) and non-K\(^+\) osmoles in patients with circulatory shock.\(^3-5\) The aim of our study was to evaluate serum osmolality in patients with INS who presented with hyponatremia and severe edema.

METHODS

Known patients with INS who were admitted to inpatient service (from November 2004 to January 2006) with generalized edema were considered for the study. Investigations performed included obtaining serum urea nitrogen (SUN), serum creatinine, electrolyte, and albumin levels, serum osmolality, hemoglobin and hematocrit levels, urinalysis, and urine protein/creatinine ratio. All investigations were performed at the time of admission before starting treatment with dietary salt restriction (<2 mEq/kg per day) and intravenous albumin and diuretics. Patients with fever, vomiting, or diarrhea or those who were already on diuretics, steroids, or immunosuppression for treatment of INS were excluded from the study.

Key Words: nephrotic syndrome, osmolal gap, hyponatremia, pre–renal failure

Abbreviations: INS, idiopathic nephrotic syndrome; CHF, congestive heart failure; SUN, serum urea nitrogen

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2554

doi:10.1542/peds.2006-2554

Accepted for publication Nov 29, 2006

Address correspondence to Tej K. Mattoo, MD, FRCP, Pediatric Nephrology, Division of Nephrology and Hypertension, 4th Floor, Carl’s Building, Children’s Hospital of Michigan, Wayne State University School of Medicine, 3901 Beaubien Blvd, Detroit, MI 48201. E-mail: tmattoo@med.wayne.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
the study. Serum sodium was measured in undiluted samples by using the principle of direct potentiometry by ion-selective electrodes (Vitrios 250 Chemistry System; Ortho-Clinical Diagnostics, Rochester, NY). Serum sodium was also measured in undiluted capillary blood samples by using a blood-gas analyzer by the same principle of direct potentiometry (ABL800 FLEX; Radiometer, Westlake, OH). Serum osmolality was measured by freezing-point depression in milliosmole-per-kilogram water. Additional tests that were performed to exclude secondary causes of nephrotic syndrome included hepatitis B and hepatitis C serology, serum complements C3 and C4, antinuclear antibody, and HIV antibody. Calculated serum osmolality was estimated by the following formula: \( 2 \times \text{Na}_e + \text{glucose}/18 + \text{SUN}/2.8 \). An osmolal gap (difference between measured and calculated osmolality) of 0 to 10 mOsm/kg water was accepted as being within normal limits.2,6,7

RESULTS
Twenty patients with INS with severe hypoproteinemia and generalized edema were admitted to our service over a period of 15 months. Of these patients, 7 (35%) were found to have low serum sodium levels as measured by direct potentiometry during routine laboratory tests. One of these patients, who had increased SUN and serum creatinine levels, was excluded from the study. The remaining 6 patients were included in the study. Their ages ranged from 2 to 11 years; there were 4 boys and 2 girls. Low serum sodium level was confirmed in each patient by capillary blood sample by using a blood-gas analyzer (Table 1).

As shown in Table 1, the measured serum osmolality was low in patients 1, 3, and 5 (reference range: 285–295 mOsm/kg water), borderline in patient 2, and within the reference range in patients 4 and 6. The calculated serum osmolality was low in all patients, and the osmolal gap was high in all patients.

DISCUSSION
Pseudohyponatremia was originally described8 when flame photometry was used for serum-electrolyte determination. Using this technique, serum samples were diluted before the actual measurement of serum electrolytes, which resulted in an artificially low serum sodium level in conditions such as hyperlipidemia, wherein the aqueous phase of plasma is reduced. However, biochemical analyzers, such as direct potentiometry, that are currently in use assay sodium concentration in the aqueous phase only and result in accurate determination of serum sodium levels.9 Given this fact, it has been suggested that the term “pseudohyponatremia” no longer be used.1 This argument is supported by recent reports on the presence of true hyponatremia in patients with multiple myeloma10 and in those who had received in-
travenous immunoglobulin infusions,\(^1\) conditions that previously were associated with pseudohyponatremia.

The exact incidence of hyponatremia in patients with INS and severe edema is not known. It is reported that serum sodium concentration is usually within the reference range in patients with INS unless it is influenced by vigorous diuretic measures or during acute water load.\(^1\)\(^2\)\(^3\) Of the 20 such patients admitted to our service, 7 (35%) had low serum sodium concentrations as measured by direct potentiometry and confirmed by blood-gas analysis.

Wang et al\(^1\) reported significantly decreased serum sodium concentrations in 5.8% of their patients with INS who presented with hypovolemic shock and in 12.5% of their patients with INS who presented with symptomatic hypovolemic episodes without hypotension. Our patients were not on diuretics, and they were clinically stable with no clinical evidence of cardiovas-
cular decompensation.

Despite true hyponatremia in all 6 of our patients, the measured serum osmolality was low in only 3 patients (patients 1, 3, and 5); the remaining 3 patients had serum osmolality that was either borderline (patient 2) or within the reference range (patients 4 and 6). The calculated osmolality was low in all patients, and all had an increased osmolar gap. Under normal circumstances, in otherwise healthy individuals, low serum sodium concentration (true hyponatremia) is associated with low measured and calculated serum osmolality and a normal osmolar gap. The presence of an increased osmolar gap in all of our patients and normal measured serum osmolality in 2 of our patients (patients 4 and 6) is intriguing.

The most common settings in which low serum sodium concentration is not associated with a decrease in serum osmolality occurs when there are additional osmoles in the extracellular fluid such as ethanol, methanol, or ethylene glycol.\(^2\)\(^3\) The coexistence of hyponatremia with an increased osmolar gap, as noted in our patients, has implications in understanding the pathophysi-ology and management of severe edema in patients with INS. We believe that the increase in the osmolal gap in patients with INS is caused by the presence of unidentified non-Na\(^+\) and non-K\(^+\) osmoles that try to maintain the intravascular volume in patients with volume contraction, as indicated in most of our patients by increased hematocrit levels and urine osmolality at the time of admission (Table 1). An increase in the osmolar gap ranging from 30 to 100 mOsm/kg water has been re-
ported in clinical and experimental models of shock,\(^3\)\(^5\) although a detailed identification of these endogenous osmoles is lacking.\(^6\) The increase in osmolar gaps in our patients (range: 10–20 mOsm/kg water) is lower than that reported in patients with shock. However, hypovolemic shock is a known complication of INS,\(^7\) and the increased osmolar gap in our patients could be indicative of intravascular volume depletion with a preshock state. Thus, patients with INS who are admitted with severe edema and hyponatremia would initially need slow volume expansion. Initial fluid restriction and diuresis in such patients could further exacerbate their volume contraction and increase the risk of pre–renal failure or ischemic acute tubular necrosis.

Our observations assume relevance when viewed in the context of recent work by Nguyen and Kurtz,\(^8\)\(^9\) who have highlighted the role of physiologic parameters, besides Na\(_e\) + K\(_e\)/TBW (Na\(_e\) = total exchangeable sodium, K\(_e\) = total exchangeable potassium, and TBW = total body water),\(^10\) that play a role in modulating serum sodium and in the generation of dysnatremias. These factors include osmotic coefficient of sodium salts, Gibbs-Donnan equilibrium, osmotic equilibrium, osmotically inactive Na\(_e\) and K\(_e\), and osmotically active non-Na\(^+\) and non-K\(^+\) osmoles.\(^8\)\(^9\) The effect of these additional parameters in the setting of hypoalbuminemia and severe edema in patients with INS needs to be studied in detail. However, on the basis of the Nguyen-Kurtz formula, it has been shown that Gibbs-Donnan equilibrium has an incremen
tal effect on serum sodium, and the presence of osmotically active non-Na\(^+\) and non-K\(^+\) osmoles in the plasma has a depressive effect on serum sodium, which may explain the presence of a relatively low serum sodium concentration and measured osmolality with the higher osmolar gap in our patients. Hypoalbuminemia in patients with INS could cause a decrease in the Gibbs-Donnan effect in these patients, contributing to hypona-
tremia.\(^9\)

CONCLUSIONS

Our results show that in children with INS and severe edema, the measurement of serum sodium by current methods indicates true hyponatremia, not pseudohy-
ponatremia. Identification of true hyponatremia is impor-tant, because the additional presence of an increased serum osmolar gap in these patients is indicative of volume contraction. These patients need slow plasma-volume expansion, because fluid restriction would increase their risk of pre–renal failure. Additional studies are needed to elucidate in detail the role of additional physi-
ologic parameters in the generation of dysnatremias in such patients.

REFERENCES

4. Inaba H, Hirasawa H, Mizuguchi T. Serum osmolality gap in


Incidence of Acute Otitis Media and Sinusitis Complicating Upper Respiratory Tract Infection: The Effect of Age

Krystal Revai, MD, MPHa, Laura A. Dobbs, MDb, Sangeeta Nair, DVM, MSa, Janak A. Patel, MDa, James J. Grady, DrPHc, Tasnee Chonmaitree, MDa,d

aDepartment of Pediatrics, bSchool of Medicine, cPreventative Medicine & Community Health, and dDepartment of Pathology, University of Texas Medical Branch, Galveston, Texas

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

ABSTRACT

Infants and young children are prone to developing upper respiratory tract infections, which often result in bacterial complications such as acute otitis media and sinusitis. We evaluated 623 upper respiratory tract infection episodes in 112 children (6–35 months of age) to determine the proportion of upper respiratory tract infection episodes that result in acute otitis media or sinusitis. Of all upper respiratory tract infections, 30% were complicated by acute otitis media and 8% were complicated by sinusitis. The rate of acute otitis media after upper respiratory tract infection declined with increasing age, whereas the rate of sinusitis after upper respiratory tract infection peaked in the second year of life. Risk for acute otitis media may be reduced substantially by avoiding frequent exposure to respiratory viruses (eg, avoidance of day care attendance) in the first year of life.

ACUTE OTITIS MEDIA (AOM) and sinusitis are 2 of the most common bacterial complications of upper respiratory tract infections (URIs) in children. It has been found that 29% to 50%1,2 of all URIs develop into AOM and 5% to 10% develop into sinusitis.3 Considering the exceedingly high incidence of URIs, the number of children affected by AOM and sinusitis comes as no surprise. On average, a child younger than 5 years of age has 2 to 7 episodes of URI per year,4,5 and a child attending day care may have up to 14 episodes per year.6 By age 3, 80% of children have had at least 1 episode of AOM,7 and ~13% have had sinusitis.8 The peak age of incidence of AOM is between 6 and 18 months, compared with 2 to 6 years for sinusitis.8 Despite the frequency of these infections and their close association with URIs, there has been no study to date that determines the age-specific incidence of AOM and sinusitis after a URI.

AOM and sinusitis are linked in several ways and may even occur concurrently. Both diseases are frequently preceded by a URI. Although AOM usually occurs between days 3 and 8 of a URI, sinusitis is not usually diagnosed until 10 to 14 days of persistent URI symptoms with no signs of improvement.9 The middle ear and sinuses have ciliated pseudostratified columnar cells that are similarly effected by viral URI. Finally, in children the 2 diseases are primarily caused by the same bacteria: Streptococcus pneumoniae, nontypeable Haemophilus influenzae, and Moraxella catarrhalis; thus, the type of antibiotic medication for treatment is similar.3,10

This study is an analysis of the age incidence of AOM and sinusitis after URI in a subgroup of children who were enrolled in an ongoing long-term study of the pathogenesis of virus-induced AOM. We calculated the overall incidence of AOM and sinusitis after a URI episode as well as the age-specific incidence and determine if age played a role in the occurrence of AOM and sinusitis after URI.

Key Words: otitis media, sinusitis, incidence, age

Abbreviations: AOM, acute otitis media; URI, upper respiratory tract infection; AAP, American Academy of Pediatrics; RR, rate ratio; CI, confidence interval; PCV7, heptavalent pneumococcal conjugate vaccine

This work was presented in part at the annual meeting of the Pediatric Academic Societies; April 29–May 2, 2006; San Francisco, CA.

Dr Dobbs’ current affiliation is Department of Pediatrics, Eastern Virginia Medical School, Norfolk, VA.

www.pediatrics.org/cgi/doi/10.1542/peds.2006-2881
doi:10.1542/peds.2006-2881

Accepted for publication Nov 28, 2006

Address correspondence to Tasnee Chonmaitree, MD, Department of Pediatrics, University of Texas Medical Branch, 301 University Blvd, Galveston, TX 77555-0371. E-mail: tchonmai@utmb.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics

e1408

REVAI et al
METHODS

Healthy children aged 6 to 35 months were enrolled from January 2003 to March 2006 in a prospective, longitudinal study of virus-induced AOM (unpublished study). Children with chronic medical problems or with an anatomic or physiologic defect of the ear or nasopharynx were not enrolled. The study was designed to capture all URI episodes that occurred during the 1-year follow-up period to study the rate and characteristics of AOM after URI. At enrollment, demographic and AOM risk-factor information was collected. Parents were asked to call the study office as soon as the child began to have URI or AOM symptoms (cough, runny nose, fever, or ear pain or tugging). Children were seen by a study physician as soon as possible after the onset of URI symptoms and then followed again a few days later (days 3–7 of the URI) for URI complications. At each physician visit, parents were asked about current URI symptoms, current medications, and history of viral illness exposure. Tympanometry was performed, and the children were examined including pneumatic otoscopy. Each URI episode was studied and monitored closely for at least 3 weeks for the development of AOM or sinusitis. AOM complicating URI was considered when the episode occurred within 21 days of the URI. AOM and sinusitis diagnoses followed established guidelines published by the American Academy of Pediatrics (AAP) and the Joint Task Force on Practice Parameters for Allergy and Immunology.

AOM was defined by acute onset of symptoms (fever, irritability, or earache), signs of inflammation of the tympanic membrane, and presence of fluid in the middle ear documented by pneumatic otoscopy and/or tympanometry. Sinusitis complicating URI was considered when children had persistent URI symptoms for >10 days without improvement or an abrupt increase in severity of symptoms, fever, or purulent nasal discharge before day 10 of illness. Children who were diagnosed with AOM or sinusitis were given antibiotic therapy consistent with the standard of care.

Parents were called twice monthly for information about current URI symptoms and the occurrence of any URI or AOM episodes since the last contact. Parents were reminded to call the study representative with any signs of URI or AOM. Any parent who could not be reached by telephone was sent a letter asking that they contact the study office. We performed an extensive chart review, including electronic medical charts, at each child’s completion of the study. Because our institution is the sole provider of pediatric health care on Galveston Island, Texas, diseases that have been diagnosed and treated in our children are very likely to be within our medical charts.

All URI, AOM, and sinusitis episodes were included in the analysis. Data were analyzed by $\chi^2$ using Stata 9.0 (Stata Corp, College Station, TX). Rate ratios (RRs) were calculated by Episheet 2001, spreadsheets for the analysis of epidemiologic data.

RESULTS

This report consists of data from 112 patients who completed the study as of October 2005. Fifty five (49%) of the patients were male, 25 (22%) were white, 34 (30%) were black, 47 (42%) were Hispanic, and 6 (5%) were of other races. The mean and median ages of the children at enrollment were 15 and 13 months, respectively. Sixty-four percent of the children were fully immunized with heptavalent pneumococcal conjugate vaccine (PCV7) according to the Advisory Committee on Immunization Practices schedule. The mean and median number of weeks of breastfeeding was 16 and 23, respectively; 37% of the children were breastfed for >2 weeks. Thirty percent of the children were enrolled in day care (18% percent of the 6- to 11-month-olds, 40% of 12- to 23-month-olds, and 44% of 24- to 35-month-olds). Twenty-nine percent of the children were exposed to cigarette smoke.

The children were followed for a total of 1231 patient-months, during which time a total of 623 URI episodes occurred; the URI episodes resulted in 188 AOM and 52 sinusitis episodes. There were 17 episodes of AOM that did not follow a URI. The study-group physicians diagnosed 77% of the AOM and sinusitis episodes. All AOM episodes that were diagnosed by a study-group physician met the diagnostic criteria according to AAP guidelines. The other 23% of the AOM episodes were diagnosed by primary care or emergency department physicians. The overall incidence of URI was 0.51 per patient-month (6.12 episodes per patient-year), of AOM was 0.15 per patient-month (2.01 episodes per patient-year), and of sinusitis was 0.04 per patient-month (0.48 episodes per patient-year). Thirty percent of all URIs were complicated by AOM and 8% were complicated by sinusitis, including 15 URI episodes (2%) that were complicated by both AOM and sinusitis (47% concomitantly and 53% sequentially; most children were given a diagnosis of AOM before sinusitis). One child had an episode of AOM followed by persistent AOM and sinusitis, and 1 child had AOM and sinusitis diagnosed at the same time but developed persistent sinusitis. The majority of AOM episodes (81%) occurred in the first 8 days of illness (range: 1–19 days). The peak days of diagnosis were on days 3 and 5 (14% and 13%, respectively). Fifty-seven percent of AOM episodes were diagnosed in the first week of the URI. Fifteen (29%) of the sinusitis episodes were diagnosed because of an abrupt increase in severity of symptoms, fever, or purulent nasal discharge before day 10 of illness; the remainder were diagnosed on the basis of persistence of symptoms beyond 10 days.

Children in the 6- to 11-month-old and 12- to 23-month-old age groups had the same incidence of URI.
(0.51 episodes per patient-month). Children in the third year of life had 0.48 URI episodes per patient-month. Children who attended day care had a URI rate of 0.59 URI episodes per patient-month, compared with 0.47 URI episodes per patient-month in children who did not attend day care (RR: 1.2; 95% confidence interval [CI]: 1.1–1.4; P = .01). AOM occurred after URI in 36% of episodes in children 6 to 11 months old (Fig 1), 29% of episodes in children in the second year of life, and 15% of episodes in children in the third year of life (P < .001 by Fisher’s exact test). Older children tended to have fewer AOM episodes after URI (P < .001 by the Cochran-Armitage trend test). Sinusitis complicated URI in 7%, 10%, and 7% of episodes in children from 6 to 11, 12 to 23, and 24 to 35 months, respectively. The rate of AOM after URI was 30% in PCV7-immunized children and 27% in PCV7-unimmunized children (RR: 1.09; 95% CI: 0.88–1.39). The rates for sinusitis were 7% and 9% of episodes in immunized and unimmunized children, respectively (RR: 0.9; 95% CI: 0.63–1.26).

DISCUSSION
It is well established that the peak age for AOM is between 6 and 18 months.7,13,14 More recent data have suggested that AOM generally occurs as a complication of viral URI.15 In this study, we clearly demonstrate that 30% of URI episodes in children result in AOM, and the disease occurs most often in children between 6 and 11 months of age, although these children are as susceptible to URI as children in the second year of life. We also found that sinusitis after URI occurs less frequently than AOM (8%), and the disease was more commonly diagnosed in children from 12 to 23 months of age.

For otitis media to occur, bacteria colonized in the nasopharynx must enter the middle ear via the eustachian tube. Normally, bacteria are prevented from entering the middle ear by the ciliated epithelium that lines the eustachian tube. Respiratory virus infection disrupts the mucociliary system and impairs the ear’s primary mechanical defense from bacterial invasion. In addition, eustachian-tube dysfunction can lead to reduced middle-ear pressure, which forces mucus, nasopharyngeal secretions, and bacteria into the middle ear; this creates an ideal milieu for bacterial superinfection.16

We found the highest incidence of AOM after URI to be between 6 and 11 months of age, and these data parallel what was found by Teele et al.7 The increased susceptibility to AOM in younger children has been postulated to be secondary to inadequate immunologic response and a shorter, straighter, and narrower Eustachian tube.10 It stands to reason that the longer children are protected from exposure to known avoidable otitis media risk factors, the later the onset of AOM and the lower their lifetime incidence.

Children who are cared for in a day care setting are exposed to many more infectious diseases, including viral URI, than children who do not attend day care.6 In our study, as in other studies,17 we show that children who attend day care have a higher burden of disease with regard to respiratory illness. Our study likely did not show a strong correlation because it was not designed to examine day care attendance. It has also been found that the later the children start group day care, the later the onset of otitis media with effusion (OME). There is a positive association between the amount of time spent in group child care and the amount of AOM.
and OME episodes. In our study, although children in their second and third years of life were more likely to attend day care and, thus more frequently exposed to viral pathogens than young infants, they were less likely to develop AOM. These data suggest that keeping the most vulnerable age group of children (eg, <1 year of age) out of the day care setting and protecting them from exposure to respiratory infections may substantially reduce AOM incidence. In children who are prone to development of AOM, parents should be advised of these risk factors, as well as other well-established risk factors such as lack of breastfeeding and cigarette-smoke exposure.

The pathophysiologic processes that occur in paranasal sinuses during a viral URI are similar to those that occur in the middle ear. The ciliated epithelium in the sinuses also loses its ability to move debris from the nasal cavity. When a child sniffs or blows his or her nose, negative pressure is formed within the sinus cavity, which draws in bacteria and debris, once again creating a model environment for bacteria to proliferate and cause sinusitis. Nevertheless, sinusitis is still a disease with a much lower incidence than AOM.

The diagnosis of sinusitis, in general, holds some degree of uncertainty, especially in children under 1 year old. Although the AAP clinical practice guidelines provide recommendations for diagnosis and treatment of sinusitis in children over 1 year old, sinusitis is sometimes diagnosed in children who are under 12 months old. In our study, there were 18 (35%) episodes of sinusitis diagnosed in infants under 12 months of age.

The majority (83%) of the sinusitis episodes in our patients were diagnosed by a study physician following the published diagnostic criteria. We found that the incidence of sinusitis was lower in the 6- to 11-month age group and in children over 24 months old when compared with children who were 12 to 23 months old, but the difference did not reach statistical significance. We postulate that the incidence of sinusitis may peak in the 12- to 23-month-old children because they are less likely to develop AOM and, thus, less likely to have received antibiotic therapy. Therefore, any low-grade bacterial infection in the sinuses would go unnoticed until cleared by the child’s own natural defense or progress to clinical sinusitis, requiring subsequent initiation of antibiotic therapy. Children in their third year of life may possibly have a lower incidence of AOM and sinusitis because they have already developed partial immunity to many microbial pathogens and subsequently do not have a strong inflammatory response to infection. This study was not powered to determine a difference in the incidence of sinusitis. Another study targeted at sinusitis may clarify these differences further.

In our study, it is possible that a number of URI episodes were not reported to us even though we were in very frequent contact with the parents. Nevertheless, the overall incidence of URI and AOM is within the ranges of those reported previously. Although the majority of URI episodes were seen by the study group, we captured another 23% of the total number of URI episodes either by parent-initiated contact, twice-monthly parental telephone calls, or chart review. Mild URI, especially in older children, may have gone unnoticed and unattended by parents; therefore, unreported URI episodes were likely to occur more often in older children than in younger ones. It is also possible that parents failed to seek medical attention for the child’s URI, or more often than for AOM or sinusitis. If missed URI episodes occurred more often than missed AOM or sinusitis episodes, then the actual incidence of AOM and sinusitis after URI could be somewhat lower than reported. Incidence according to age group should still be proportionate or even lower than reported in older age groups, which would widen the difference between the 3 groups. URI, AOM, and sinusitis diagnosed by physicians other than those in the study group were unlikely to be missed.

CONCLUSIONS

We found that children from 6 to 11 months of age were at highest risk for developing AOM after URI. Older children were more likely to attend day care yet developed fewer episodes of AOM and sinusitis. Delaying entry into group day care until the second year of life could help reduce the incidence of AOM in infants and young children.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grants R01 DC005841 and DC 005841-02S1. The study was conducted at the General Clinical Research Center at the University of Texas Medical Branch, which is funded by National Center for Research Resources (National Institutes of Health, US Public Health Service) grant M01 RR 00073.

We thank M. Lizette Rangel, Kyralessa B. Ramirez, Liliana Najera, and Rafael Serna for assistance with study subjects.

REFERENCES

Electronic Prescribing Systems in Pediatrics: The Rationale and Functionality Requirements

Robert S. Gerstle, MD, Christoph U. Lehmann, MD, and the Council on Clinical Information Technology

ABSTRACT
This technical report discusses electronic prescribing systems and their limitations and potential benefits, particularly to the pediatrician in the ambulatory setting. In the report we acknowledge the benefits of integrating these systems with electronic health records and practice-management systems and recommend that the adoption of electronic prescribing systems be done in the context of ultimately moving toward an electronic health record. This technical report supports the accompanying American Academy of Pediatrics policy-statement recommendations on the adoption of electronic prescribing systems by pediatricians.

INTRODUCTION
Electronic prescribing (e-prescribing) systems are computer applications designed for use by clinicians to generate paper or electronic medication prescriptions. They offer the clinician and the patient the promise of safer prescribing and improved office efficiencies, 2 major drivers for the adoption of such systems. Many organizations, notably the Institute of Medicine (IOM), the Institute for Safe Medical Practice (ISMP), and the Leapfrog Group, involved in efforts to improve medical quality and reduce medical errors, have endorsed e-prescribing systems as a major tool in reducing medical errors. The American Academy of Pediatrics has recognized that within the hospital, computerized physician order entry (CPOE) can prevent medication errors.

One significant cause of medication errors has been the misinterpretation of physician handwriting. The National Hospital Ambulatory Medical Care Survey suggests that the prescription illegibility rate may be 1% to 2%. Illegible orders may account for 30% of errors.

In their simplest implementation, e-prescribing systems provide printer-generated, easily readable prescriptions that are less likely to be misinterpreted or misread by a pharmacist. The most completely envisioned e-prescribing systems provide extensive decision support to the clinician and offer additional office efficiencies designed to streamline the prescribing and renewal process, facilitate insurance formulary adherence, and improve prescribing safety. This technical report will review the current state of e-prescribing to provide the clinician with an understanding of the subject and to better guide clinician decision-making in the adoption of this technology in the office setting.

CPOE is the component of the clinical information system that allows prescrib-
ers to enter orders (for medications and/or clinical procedures) directly into a computer for electronic processing and transmission to appropriate departments and/or individuals for completion. The most obvious advantages of CPOE include the immediate transmission of orders without further transcription or delay to (usually) hospital departments, which eliminates transcription errors by nurses and speeds the health care–delivery process. Most research on CPOE has been performed in the inpatient setting. Analysis of research done on inpatient prescribing by the IOM resulted in the conclusion that 44 000 to 98 000 inpatient deaths could be attributed to avoidable medical errors, and many of these deaths could potentially be prevented by CPOE. In an analysis at 1 academic medical center, 64.4% of errors were rated as likely to be prevented with CPOE (including 43% of the potentially harmful errors).

Less is known about the error rates and resulting morbidity and costs to the medical system that result from prescribing errors in the outpatient setting. In 1 study, the use of e-prescribing in the emergency department demonstrated a reduction in prescriptions that contained errors (odds ratio [OR]: 0.31) or required pharmacist clarification (OR: 0.19). The Center for Information Technology Leadership projects that use of e-prescribing would prevent more than 3 million adverse drug events (ADEs) and prevent 190 000 hospitalizations yearly. Potential financial benefits from the universal adoption of e-prescribing include a savings of $27 billion/year, including $2 billion/year from the reduction in ADEs. As noted previously, knowledge of the specific health, safety, and financial impact of drug errors and e-prescribing in the pediatric outpatient setting is considerably less well understood.

Despite these limitations, there is a groundswell for the adoption of e-prescribing by the medical community. Organizations such as the IOM, ISMP, and Leapfrog Group have strongly endorsed e-prescribing: the Centers for Medicare and Medicaid Services will require e-prescribing as part of its Medicare prescription drug program; and insurers are looking to these systems as a way to avoid costly medical errors and to “enforce” compliance with tiered formularies to drive down their pharmacy costs. Thus, a number of commercial insurers are now underwriting the upfront costs for physicians who are willing to adopt this new technology. Currently, 5% to 18% of physicians are using some type of e-prescribing system in their offices, with this number expected to grow rapidly over the next 3 to 5 years.

**BENEFITS**

E-prescribing has potential benefits for public health, the patient, the insurer/pharmacy benefits manager, and the physician. Specific benefits to each are noted below.

**Benefits to public health include:**
- reduction in medical errors and associated costs to society;
- reduction in drug diversion (forgery);
- improved patient care and improved health outcomes; and
- improved efficiency and reduced costs associated with prescribing.

**Benefits to patients include:**
- reduced chance for medication misadventures; and
- improved patient satisfaction.

**Benefits to pharmacists, pharmacy benefit managers, and insurers include:**
- workflow efficiencies;
- improved compliance with formulary prescribing, with attendant reduction in drug costs; and
- reduction in costs attributable to preventable ADEs.

**Benefits to the physician include:**
- improved office efficiencies for medication renewals;
- incentive reimbursement for compliance with formulary programs; and
- improved record keeping and documentation.

**CPOE AND E-PRESCRIBING SYSTEMS REDUCE MEDICAL ERRORS AND IMPROVE OUTCOMES**

**Hospital-Based Studies**

Accurately determining the frequency of medication errors is limited by findings that medical errors are significantly underreported in the incident-reporting systems used by many hospitals, especially when errors did not reach the patient (near-miss events). Despite this limitation, however, 2.4% of hospitalized patients develop a clinically significant ADE during their hospitalization. Medication errors are one of the most common medical errors and the most frequent cause of adverse events, accounting for 19% to 20% of all adverse events. In pediatric patients, the most common type of medication error is a dosing error at the ordering stage.

Preventable ADEs (eg, ordering and administering an incorrect medication or dosage) are more common than nonpreventable ADEs (eg, a newly developed drug allergy). In 1995, Leape et al described poor dissemination of drug knowledge (29%) and inadequate availability of patient information (18%), as well as the results of laboratory tests, as the most common causes of ADEs.

An analysis of significant ADEs concluded that 52% of the cases were preventable; of these events, 50% could have been prevented by a pharmacist.
E-prescribing tools have been used successfully to reduce errors in the prescribing process. Use in the emergency department demonstrated a reduction in prescriptions that contained errors (OR: 0.31) or required pharmacist clarification (OR: 0.19).11

Decision support provides knowledge, information, and data to the provider to optimize the selection and specification of medication for ordering. Such knowledge may be provided at the point of care through updated formularies for handheld devices and online guidelines and tools for specific domains such as infectious diseases.21,22 A computer-assisted antibiotic-management program was able to reduce ADEs and reduce total costs in 1 study23 and reduce length of stay and antibiotic use in another.24 E-prescribing tools have been used successfully to prevent the prescription of gender-specific drugs for patients of the opposite gender.10

Pediatric Studies
Children are at increased risk of certain specific types of ADEs. Pharmacologic factors, including age-based variability in absorption, metabolism, and excretion of drugs as compared with adults, pose special vulnerabilities to the adverse effects of overdosing (often by an order of magnitude). Physiologic factors, such as the nearly universal need for weight or body surface area considerations in dosing and recognition of the variability of organ development, also make the medication process for pediatric patients more prone to dosage errors than for adults.4,25

Error rates for children seem to be inversely related to the weight of the patient, with infants in the NICU being most likely to experience medication errors and potential ADEs.26,27 A study of errors for preterm neonates before discharge demonstrated a linear increase in medical error rates that was inversely related to birth weight, although the overall rate of errors was lower in comparison with children and adults.28 Medical errors in pediatric patients are more likely to be caused by calculation or dose errors than in adults. Medications in adults that require weight-based calculation (the norm for pediatric patients) were found to be more error prone.29,30

Process factors, including the need for individualized dilution of stock medications and fluids (because of weight and body surface area considerations), place children at increased risk of medication errors in comparison with adults. Location-specific factors, such as the fast pace and high complexity in ICUs, are associated with a sevenfold risk of medication errors.31

Medication errors may occur at any step in the process, from ordering (56%) to transcription (6%), dispensing (4%), and administration (34%).12-34 Orders by prescribers are the most error-prone steps in the medication process, with the wrong dosage being the most common type of error.35-37 These errors may or may not be caught by subsequent checks, such as during dispensing and administration.38 In pediatric inpatients, almost three quarters of all medication errors were discovered in the ordering stage.27 In academic pediatric critical care settings, prescribing error rates of 11% to 30% were observed, compared with 6% of prescription errors in an internal medicine setting.10,39,40

The evidence for error reduction in pediatric patients using computerized systems is not yet robust. A recent Cochrane review41 concluded that limited data from randomized trials exist on which to assess the effects of clinical decision-support systems in neonatal care. However, some evidence that highlights the benefits of CPOE in the neonatal population is emerging. Cordero et al42 reported that implementation of CPOE resulted in a significant reduction in medication turnaround times and medication errors for selected drugs (gentimycin) and a decrease in ancillary service (radiology) response time. CPOE eliminates illegibility. Vanderbilt University’s WizOrder system reduced the rate of errors caused by illegible pediatric intensive care orders from 1 error per 100 orders to zero.39 CPOE with calculators and point-of-care decision support has also been used successfully to drastically reduce provider errors in the ordering of total parenteral nutrition43 and continuous infusions.44 An ambulatory study from Singapore showed that automated calculation reduced pediatric prescribing errors from 28.2% to 12.6%.45 CPOE has also been shown to reduce errors in the ordering of chemotherapy agents in pediatric patients.46

Ambulatory Care Studies
ADEs are common in ambulatory care, and many are preventable or ameliorable.34 E-prescribing tools in outpatient settings have been used successfully to lower prescription costs through electronic, evidence-based decision support during the prescribing process.47 Computerized prescription systems have been shown in randomized trials to improve the quality of anticoagulation.48

LIMITATIONS OF CPOE AND E-PRESCRIBING IN REDUCING ERRORS
Computerized ordering and prescription tools have been advertised as means to reduce the frequency of ADEs.49 However, evidence exists that computerized systems cannot prevent all errors or ADEs and may, in some situations, be responsible for new types of errors.50,51 A recent study of a 110-bed computerized Veterans Administration hospital found 70 ADEs per 100 patient-days (significantly higher than previously reported).14 The authors suggested that the legible and accessible electronic records may have facilitated the increased identification of ADEs. Of note in the Veterans Administration study is the finding that errors in ordering (74%) accounted for a larger percentage of errors than previously reported (56%).35 At the same time, errors
during transcription and administration were reduced. The authors speculated that these findings were a direct result of the system design—a computerized system that eliminates need for transcription and ensures legibility but lacks decision support for drug selection and dosing—and will “redistribute” error frequencies. In other words, unless an electronic system is designed to prevent errors at the ordering stage, it will not prevent these errors; on the contrary, it will increase the speed at which these errors can be committed and executed. Another recent study from the University of Pennsylvania that evaluated an older CPOE system with very limited decision support received significant media attention when the authors concluded that a leading CPOE system often facilitated medication-error risks. Although there was no comparison between manual ordering and CPOE, the authors emphasized that electronic systems will do just what they are designed to do. If they are not designed to provide decision support, they will not do so. In contrast, a study of the WizOrder CPOE system at Vanderbilt University demonstrated that a CPOE system that includes sophisticated decision support at the point of order entry may reduce medication-prescribing errors by 99.4% and rule violations (deviations from ordering policy) by 97.9%. A recent study in a PICU showed that implementation of a CPOE system, even in the early months after implementation, was not associated with an increase in mortality.

In addition to incomplete design, computerized prescription-writing tools are limited to the content of the program. For example, they may have a limited drug inventory or lack dose-range information and may, especially when used on handheld devices, pose usability problems. Fernando et al, who used simulated test cases, found that computing systems currently in use in approximately three quarters of general practices in the United Kingdom have clinically important safety deficiencies.

The Rand Electronic Prescribing Expert Advisory Panel has provided 60 capability recommendations for ambulatory prescribing systems. However, a recent study of 10 commercially available e-prescribing systems demonstrated that only 50% (range: 26%–64%) of these recommendations were implemented. It must be recognized that electronic systems are only as good as they are designed and implemented. A system that strives to provide legibility and accessibility only will improve the error rate in transcription and administration processes but not in ordering processes. For CPOE and e-prescribing systems to reduce provider ordering errors, they must be integrated with sophisticated clinical decision-support capabilities.

**E-PRESCRIBING SYSTEMS**

An e-prescribing system, at its simplest, is a computer application that allows physicians to print out prescriptions (a word processor would fulfill this definition). The advanced vision of an e-prescribing system, however, is that of an application that facilitates the rapid and efficient generation of prescriptions (including electronic renewals); includes a knowledge base with drug information relevant to the prescribing process; performs all necessary calculations automatically and accurately; checks for prescription completeness, drug contraindications, drug interactions, allergies, and medical conditions that affect prescribing; verifies appropriateness of dose on the basis of patient age, weight, gender, and medical conditions (eg, renal insufficiency); checks insurance formulary preferences; and then transmits the prescription electronically to the pharmacy. An ideal closed-loop system would also receive data back from the pharmacy to confirm that the patient has filled the prescription.

Although generating printed prescriptions would be expected to reduce medical errors related to handwriting, it is the integrated clinical decision support, automatic calculations, and electronic transmittal functions that are likely to have the greatest effect on physician prescribing habits and to improve the safety and efficiency of the prescribing process. Merely printing out prescriptions can be done on a word processor; a fully functional e-prescribing application requires significant domain knowledge often contained in database tables and is most effective when integrated (bidirectional data transfer) with office practice–management and/or electronic health record (EHR) systems. Higher levels of systems are associated with higher startup costs and complexity and are generally associated with higher benefits. These levels of e-prescribing have been described.

**Electronic Drug Reference Only; No Prescription-Writing Capability**

This functionality is supplied by commercially available software programs, many of which are designed for mobile personal digital assistants, that allow access to drug dosages, contraindications, adverse effects, and drug interactions. It is important that reference data be kept current and updated at least monthly.

**Stand-Alone Prescription Writer With No Medication History or Supporting Data**

In addition to providing electronic drug references, a stand-alone writer provides computerized printing of prescriptions that are then given to the patient or manually faxed to the patient’s pharmacy.

**Reference Data and Prescription Writer With the Addition of Basic Supporting Data Such as Allergies, Demographics, Past Prescriptions and Formulary Information, Which Can Be Used to Generate Alerts**

This functionality allows the application to incorporate clinical decision support (including drug-allergy, drug-
Medication Management: Long-term Tracking and Monitoring of Each Patient’s Active Medications

This level contains the previous functionality and maintains a database of the patient’s previous prescriptions and prescription renewals. These applications typically monitor for drug-drug interactions automatically. These systems should also allow for the manual entry of other medications the patient is taking. Less commonly available but useful is the ability to enter alternative and nonprescription medications. Some vendors offer the ability to check for drug interactions with alternative medications.

Unidirectional Connectivity From Practices to Pharmacies, Payers, Pharmacy Benefit Managers, or Clearinghouses

This type of system typically provides the previous functionality and allows for the electronic transmission of prescriptions to pharmacies and often includes subscriptions to electronic versions of insurance formularies to identify preferred and tiered drugs and alert for noncovered medications. This requires that patient insurance information be entered into the e-prescribing system or transferred from practice-management systems.

Integration With a More Complete EHR

Systems integrated with an EHR allow for a wider range of clinical decision support without the need to manually reenter data into the e-prescribing system. They also automatically update the patient’s current medication list within the EHR.

Bidirectional Connectivity Between Physicians, Pharmacies, Payers, Pharmacy-Benefit–Management Programs, or Clearinghouses

This functionality is not generally available in systems today but is in the planning stages. It would allow for feedback of prescription information from the pharmacy to the clinician, such as confirming that the prescription has been received, has been given to the patient, or is overdue for refilling, thus enabling compliance monitoring by the physician and improved medical management.

Bidirectional functionality could allow physicians to receive up-to-date information from other physicians’ prescribing systems or from the pharmacy on their patient’s prescriptions, such as information on prescriptions prescribed by another physician or in another care setting. Up-to-date medication lists are essential for accurate drug-interaction checking.

BARRIERS AND POTENTIAL SOLUTIONS TO E-PRESCRIBING ADOPTION

Failure to Recognize Current System Deficiencies

Often, physicians do not perceive that they make prescription errors or have illegible handwriting. They do not perceive themselves as part of the drug-error problem and are often reluctant to change their practices.56

Technology Barriers (Equipment Setup and Maintenance) in the Office Setting

The lack of access to a broadband Internet connection especially affects smaller and more rural practices. Those practices may also suffer from lack of access to the technological support they need.

Implementation, Training, and Maintenance Cost

Establishing e-prescribing in the office is not a 1-time expense. Licensing and maintenance-agreement costs are ongoing and are generally not offset by a reduction in other office expenses. E-prescribing may not be more time-efficient than handwriting prescriptions and could affect office productivity, particularly in the initial implementation phase. Additional manual data-entry requirements may also reduce efficiency. Potential office cost benefits may result from improvements in the medication-renewal process, reduction in manual chart pulls, and compliance with insurer incentive programs for using generic or preferred drugs. However, even these efficiencies may not always allow for a reduction of office staff, which would equate to decreased office salary expense.

Beneficiary-Payer Discrepancy (Misaligned Incentives)

Although providers must carry the bulk of the investment for e-prescribing systems, benefits from automation are more likely to accrue primarily to others such as insurers, pharmacists, and patients. Patients appreciate the potential of having prescriptions ready to be picked up without a wait. Without incentives to physicians to invest in e-prescribing systems, the adoption of this technology by physicians is likely to be slow.12

Interface/Integration Costs

It is likely that physicians in the future will want to implement a complete EHR system in their offices. Transferring data from the e-prescribing system or integrating the e-prescribing system with an EHR system can be difficult and expensive. Physicians expecting to move
to an EHR system in the near-to-medium future will need to build in a transition strategy to avoid duplicate data-entry costs.

Existing Legacy Systems
Although most pharmacies can accept facsimile transmissions, many still do not have the ability to electronically accept transmitted prescriptions into their electronic pharmacy systems and, thus, miss out on potential benefits. Even where pharmacies can accept the electronic transmissions or facsimiles, pharmacy workflow may be such that the drug is not dispensed before the patient arrives to pick up the medication.

Legal and Regulatory Barriers
In regard to e-prescribing, nonuniform state regulations and lack of federal standardization (preempting state regulations) place an additional burden. In particular, regulations on controlled substances that mandate triplicate prescriptions and special forms may significantly limit the use of e-prescribing. Pediatricians commonly prescribe schedule II controlled substances (eg, stimulants for attention-deficit/hyperactivity disorder) and might, therefore, not be able to fully benefit from e-prescribing. The Drug Enforcement Agency currently allows e-prescribing of controlled drugs in category III to VI and is considering issuing digital certificates and using public-key infrastructure encryption to allow for digital prescription for category II drugs.90 Until state and federal legislatures generate a uniform regulatory approach, e-prescribing systems are required to fulfill all mandates or to be tailored for use in specific states.

Negative Past User Experiences
Poorly designed systems have been available for several years and may have given the market a bad reputation. Current systems are evolving to be more user friendly, intuitive, and customizable. Nonetheless, any potential user of an e-prescribing system should test a system that is in actual use before investing the significant resources required. Testing should include some of the more complex prescriptions used in practice, such as prednisone tapers, drugs that require additional information or diagnosis to be included on the prescription, drugs with as-needed indications, and over-the-counter (OTC) drugs (covered by some insurers with prescription). Physicians must be aware of the fluidity of the vendor market and must appreciate the risk involved in the bankruptcy, sale, or merger of their e-prescribing system vendor and the possible subsequent discontinuation of software maintenance and technical support for their system.

Lack of Standards
Several transmission standards for prescriptions exist; however, a move toward consolidation of standards has been underway, and the recent Centers for Medicare and Medicaid Services move to establish electronic standards under the Medicare Part D prescription drug plan will help to establish a de facto standard. However, these rules, regulations, and standards only apply to prescription drugs. E-prescribing systems may not be set up to handle prescribing or provide decision support for OTC and alternative medications.

OPTIONS AND FUNCTIONAL REQUIREMENTS
There are many ways to implement e-prescribing systems; solutions that work for one practice may not work for another. Careful consideration of the risks and benefits of implementation alternatives is necessary. It is critical that physicians considering the use of e-prescribing be aware of the risks including costs (purchase, training, workload, and maintenance) as well as legal and regulatory requirements in their state.

All systems that are capable of electronically transmitting prescriptions share certain characteristics such as a need for connectivity. Most will require dedicated telephone lines or broadband Internet connectivity, a potential problem in more remote areas. All of them will require a computer, modem to connect to the telephone (usually dedicated digital subscription line) or Internet (via digital subscription line or cable), and likely a router. The need for connectivity may introduce a single point of failure in which a malfunction may render the whole system inoperable, particularly for those applications that run as an application service-provider (ASP) system.

If more than 1 computer or device in the office is to be used for prescribing, a computer network or a way to synchronize information will be necessary. The input equipment can be a computer (desktop, laptop, tablet computer), wireless handheld device, or personal digital assistant. The network can be wired, wireless, or a combination of both. A wireless router and perhaps several wireless antennas will be required for wireless networks. When local printing of prescriptions is planned, printing may be centralized within the office, or multiple printers may be necessary, possibly convenient to each examination room. If integrated electronic facsimile transmission is not part of the e-prescribing system, a fax machine for manual transmission may be required. Technical help is usually required to wire and set up networks.

Systems can be administered from off-site centralized locations through an ASP, where everything is “done for you” to manage the system, or maintained on-site in the office. ASP systems use secure connections to allow office access devices to be logged into the remote system. With the ASP model, patient data are typically stored off-site, but in either case regular data backups need to be ensured.

Computer interfaces to transfer data from other office
systems (eg, practice management or scheduling) can initially be expensive but may save providers time by eliminating the need to enter demographic data and by keeping data current and synchronized across multiple systems with single data entry. There are usually additional costs associated with purchasing and integrating various databases, including insurance company drug formularies, prescribing-information updates, decision-support data, and pharmacy lists with up-to-date fax numbers.

A number of insurers have been paying the implementation costs for selected e-prescribing systems, at least for their high-volume prescribers. Insurance companies expect to realize benefit from providing e-prescribing systems to providers by steering physicians to the selection of specific preferred drugs and the reduction in ADEs to insured clients. Some medical societies have negotiated “discounted” deals with preferred e-prescribing providers. It is estimated that the direct cost of implementing a stand-alone system can be under $2000 per physician and, in some cases, considerably less. However, the cost of implementing an integrated EHR with an e-prescribing component is considerably higher. Providers will need to determine if an integrated EHR solution, although more expensive and difficult to implement, will have other offsetting benefits that would justify the additional work and expense.

Physicians must be aware that the market for these applications is somewhat “fluid” as vendors come, go, and merge with others. Choosing a product from a well-established, financially stable vendor, although no guarantee of sustainability, will help to ensure longer-term product support. Thus, it is important that an “exit strategy” be in place, including provision for recovery of the data contained in the system (in a standard nonproprietary database format). Ongoing costs are for equipment depreciation and replacement, renting telephone or Internet access, licensing and maintenance-agreement costs, and staff training and can vary considerably.

An expert consensus panel recently published recommendations for comparing e-prescribing systems. They categorized the functionality to be assessed into the categories summarized below. They noted that no current system met their recommended criteria fully.

**Patient Identification**
Patient name, gender, and date of birth or age must be visible throughout the ordering process. These data should be imported from other systems when available, or the system should allow for manual entry when not imported. Duplicate records created under separate identities for one individual should be able to be reconciled and merged. The ability to perform patient searches using combinations of date of birth, gender, and partial name helps to positively locate and identify patients.

**Access to Patient Historical Data**
The system and clinicians should have access to external sources of data (hospitals, pharmacies, laboratories, EHRs) and be able to review all patient prescriptions, not just those written by the current provider, as well as OTC and alternative medications taken by the patient. The ability to manually enter additional patient medications is required. Current and past medication-prescription details should be viewable by class, with start and stop dates. When a medication is discontinued or changed, a notification should be sent to the original prescriber. When a diagnosis is entered, a list of medications by diagnosis should be viewable.

**Medication Selection**
When the e-prescribing system is presented with a patient diagnosis, a customizable selection of appropriate medications should be presented to the user. Prepopulated lists or dynamically maintained lists of common medications based on prescribing frequency can speed medication ordering. Medication options should not be influenced by vendor or insurer promotional considerations.

When the system displays a preferred drug, the rationale for that preference should be immediately viewable, and contraindicated medications (based on allergy or drug interactions) should not be viewable as a preferred medication choice. Prescribing by medication name should override restricted-medication menus. The system should provide formulary status and cost to the patient for medication options on the basis of insurance and any benefit or prescribing caps. Clinical summary data useful to the selection process should be easily accessible.

The user should be able to easily select the drug form and available strengths for each medication. Dosing recommendations based on calculations of body size (weight or body surface area) and age, when appropriate, should be available. When appropriate, adjustments calculated and based on renal and liver function should be made. The ability to default to generic drug name on prescriptions should be available (unless specifically overridden) to aid in insurance prescribing compliance programs and reduced costs to patients.

**Alerts and Other Messages to Prescribers**
Dose-range checks based on dose, dose per weight, daily dose, daily dose per weight, and lifetime dose-checking alerts should be available. Prescribers should be alerted when there is a contraindication or precaution based on allergy, drug interaction, medical condition, or laboratory results. The system should send a reminder when a medication is indicated in a particular instance (immunization due or medication based on diagnosis, laboratory results, and peer-reviewed recommendations). Messages should provide a clear rationale for any rec-
ommendations. Messages not based on patient safety concerns should be suppressible to avoid “alert fatigue.” A user should be able to prioritize safety alerts and set a threshold that allows only alerts of a certain level/priority to result in a message to a provider while other alerts are suppressed. Providers should be able to override alerts with reasonable justification, and the system should track all changes and their justification. Prescribers should be able to flag or correct (update) incorrect information.

Patient Education
Patient medication-information sheets, written at an appropriate level for patients and their parents, and patient medication lists should be printable. Information sheets should be editable or customizable (and then saved for reuse) for pediatric indications (eg, \( \beta_2 \) blockers used for migraine control, not heart disease; digoxin for arrhythmia control, not congestive heart failure). Patient education materials should be able to be edited or personalized and be available in other languages in addition to English.

Data Transmission and Storage
Prescriptions should be able to be transmitted to the patient’s pharmacy of choice. Transmission should conform to current Health Level 7 (HL7) and National Council for Prescription Drug Programs standards. Physicians should receive transmission and dispensing receipts and should be notified of any transmission failures.

When prescriptions are printed locally and given to the patient, prescription abbreviations should be avoided, and the prescriptions should be consistent with best-practice recommendations (eg, Joint Commission on Accreditation of Healthcare Organizations or ISMP).59,60

Monitoring and Renewals
The prescriber should be notified electronically when a prescription or refill is not dispensed within a provider-specified time period. Ideally, the system should alert the clinician to place orders for manufacturer-recommended laboratory monitoring and alert the physician when laboratory results require action. Prescriptions and renewals entered should be clearly attributable to the person who enters the order.

Many e-prescribing systems have a messaging ability integrated into the system so that nurses and clerical staff can access the system and forward renewal requests that come in directly from patients. Office processes and staff training needs require attention.

Transparency and Accountability
The system should clearly display any sponsorships or relationships that could represent conflicts of interest and any sources and methods used to develop clinical decision-support rules and messages.

Prescriber-Level Feedback
Prescribers should be able to review profiles of their own prescribing patterns and history of overriding alerts.

Security and Confidentiality
Systems must be compliant with the Health Insurance Portability and Accountability Act (HIPAA). Access to protected health information should be auditable. Each user should have a unique sign-on and password and role-based access privileges. The system should support data integrity checking of stored and transmitted data. Provisions for the routine backup of data and secure storage must be considered. Firewalls may be needed to protect systems, and antivirus software must be current if the network is not dedicated to the e-prescribing application. Access-management processes must be in place (eg, to revoke access when an employee is terminated).

CONCLUSIONS
Ultimately, the issue for consideration is “not if, but when.” A uniform system for providing incentives and removing barriers to the adoption of e-prescribing systems by physicians who wish to use these systems will likely be needed to accelerate the migration to e-prescribing. The expenditure of time and money to implement a well-designed e-prescribing system has the potential, in the long run, to benefit society, the patient, the insurer, and the physician.

COUNCIL ON CLINICAL INFORMATION TECHNOLOGY, 2006–2007
Mark M. Simonian, MD, Chairperson
Joseph H. Schneider, MD, MBA, Vice Chairperson
Kristin A. Benson, MD
Donna M. D’Alessandro, MD
Mark A. Del Beccaro, MD
Willa H. Drummond, MD, MS
George R. Kim, MD
Michael Leu, MD
Gregg C. Lund, DO
Eugenia Marcus, MD
Alan E. Zuckerman, MD

PAST COUNCIL EXECUTIVE COMMITTEE MEMBERS
S. Andrew Spooner, MD, MS, Immediate Past Chairperson
Robert S. Gerstle, MD
Kevin B. Johnson, MD, MS
Christoph U. Lehmann, MD

STAFF
Rebecca Marshall
REFERENCES


29. Yasir T, Briceiland L, Stein DS. Factors related to errors in medication prescribing. JAMA. 1997;277:312–317


60. Institute for Safe Medication Practices. Special issue: do not use these dangerous abbreviations or dose designations. Available at: www.ismp.org/Newsletters/acuteCare/articles/20030220.2.asp. Accessed May 22, 2006
Use of Sleep Medications in Hospitalized Pediatric Patients

Lisa J. Meltzer, PhD; Jodi A. Mindell, PhD; Judith A. Owens, MD; Kelly C. Byars, PsyD

Objective. Little is known about the medications prescribed for sleep in hospitalized children. The aims of this study were to (1) determine the percentage of hospitalized children who receive medication for sleep disturbances, (2) determine what medications are prescribed for sleep difficulties, and (3) examine medical and demographic variables related to medications prescribed during hospitalization.

Patients and Methods. A chart review was conducted for all inpatients at 3 pediatric hospitals across 26 randomly selected days in 2004. Demographic, medical, and medication data were collected on 9440 patients. The sample was 54.5% male, had a mean age of 7.0 years, and was 63% white. Almost 19% of the patients had at least 1 psychiatric diagnosis.

Results. Overall, 6.0% of all hospitalized children (3% of all medically hospitalized children, excluding children with a psychiatric diagnosis) were prescribed medications for sleep, with antihistamines the most frequently prescribed medication (36.6%), followed by benzodiazepines (19.4%); hypnotic agents were the least frequently prescribed (2.2%). Significant differences were found in both the frequency of sleep-medication prescriptions and the types of medications used across hospitals, as well as for age, length of hospitalization, and service that the child was discharged from. Children with a psychiatric diagnosis were more likely to receive a sleep medication, with 22% of children on a psychiatric service receiving a sleep-related medication.

Conclusions. Approximately 3% to 6% of children are treated pharmacologically with a broad array of sleep medications in hospital settings. Prescribed practices vary by hospital, medical service, child age, and diagnosis. The results from this study indicate that medications are being prescribed for sleep in hospitalized children, especially in children with psychiatric diagnoses. However, given that there are neither Food and Drug Administration–approved sleep medications for children nor clinical consensus guidelines regarding their use, clinical trials, practice guidelines, and additional research are clearly needed.
SLEEP DISTURBANCES ARE one of the most common behavioral problems experienced by children and adolescents. They are especially prevalent in special populations, including hospitalized children, and few data exist on the treatment of sleep disturbances in this population. A task force on pharmacotherapy for pediatric insomnia sponsored by the American Academy of Sleep Medicine recommended that although nonpharmacological treatments such as sleep hygiene should be the first line of treatment for pediatric insomnia, hypnotic agent use is indicated when the child’s insomnia “occurs in the setting of medical illness with associated issues, including pain control, concomitant medication, and/or hospitalization.” However, there are no medications currently approved for pediatric insomnia by the Food and Drug Administration (FDA), thus it is unclear what medications physicians are using to address sleep problems in hospitalized children.

Several surveys of general pediatricians have found that a wide variety of sleep medications are prescribed commonly in nonhospitalized children. A recent study found that >75% of 671 community-based pediatric practitioners had recommended nonprescription medications and >50% had prescribed a medication for “insomnia” (defined as significant difficulty falling or staying asleep) in the past 6 months. Alpha agonists were the most frequently prescribed sleep medications (31%), followed by prescription antihistamines (29%), antidepressants (16.4%), and benzodiazepines (12%). Chloral hydrate (12%) and nonbenzodiazepine hypnotic agents (8%) were also prescribed. There were significant differences in types of medications used depending on the age group of the child.

In addition, a recent survey of 1271 practicing child and adolescent psychiatrists found that insomnia was endorsed as a major problem seen in 28% to 32% of school-aged and adolescent patients, respectively (J.A.O., C. L. Rosen, MD, J.A.M., and L. H. Kirchner, MD, unpublished data, 2006). Medication use by child and adolescent psychiatrists for treating insomnia was very high. For example, for children with attention-deficit/hyperactivity disorder (ADHD), respondents prescribed a variety of medications to treat insomnia in a typical month, including alpha agonists (clonidine; 81%), sedating antidepressants (71%), trazodone (60%), atypical antipsychotic agents (34%), and nonbenzodiazepine hypnotic agents (18%). For children with insomnia associated with mental retardation and developmental delay (autism and pervasive developmental delay).

In another study exploring prescriptions for sleep problems, 20% of children had received at least 1 dose of a sleep medication. This study involved >38 000 Medicaid recipients in Michigan. There was a wide variation in prescribing practices by region in the state and by individual practitioner. Finally, a recent analysis, which reviewed prescription drug claims of 2.4 million Americans between 2000 and 2004, reported that the number of children and adolescents between the ages of 10 and 19 years using sleeping medications rose from 0.16% to 0.3%, an increase of 85%. Across the entire pediatric age range (0–19 years), a total of 39% of the patients who were prescribed sleep medication were also on another psychotropic medication.

It is expected that hospitalized children are at risk for pediatric insomnia because of both environmental factors related to hospitalization and medically related issues, such as pain and anxiety. Few studies have investigated sleep problems in this population, although several studies have been performed in adult hospitalized patients. Studies of adult patients have found reduced sleep time and sleep efficiency, as well as increased awakenings and more daytime sleep. Up to one third of adult patients report insomnia, and between 50% and 75% of adult hospitalized patients are prescribed a sedative or hypnotic agent to improve sleep. Less is known about pediatric sleep disturbances while hospitalized, although sleep problems are common. Hagemann found that hospitalized children 3 to 8 years old lose 20% to 25% of their expected sleep time, primarily because of prolonged sleep latencies. Other studies by White and colleagues noted significantly later bedtimes in children when hospitalized. No studies, however, have investigated the use of medications to treat sleep disturbances in hospitalized pediatric patients.

Thus, the purpose of this descriptive study was to examine the use of sleep medications in hospitalized children. The aims of this study were (1) to determine the percentage of hospitalized children who receive medication for sleep disturbances, (2) to determine what medications are prescribed for sleep difficulties, and (3) to examine medical and demographic variables related to medications used for sleep difficulties in pediatric hospitals.

METHODOLOGY

Participants and Procedure

Twenty-six days in 2004 were selected randomly for a chart review of patients who were hospitalized on those dates at each of 3 children’s hospitals (Children’s Hospital of Philadelphia (CHOP; 381 beds), Cincinnati Children’s Hospital Medical Center (CHMCC; 324 beds), and Hasbro Children’s Hospital (HCH; 87 beds). The range of
dates controlled for biases that may have resulted from time of year, day of week, and inpatient staff coverage. Electronic medical records were reviewed at CHOP and CCHMC, whereas HCH reviewed paper medical records. Because of an absence of data during a changeover of the medical records system at CHOP, only 25 days were included, and because of the labor intensity of reviewing records by hand, only 17 days at HCH were included. Patients who were admitted for >1 of the selected dates had their data included in the database only for the first date that they appeared. This resulted in 9440 patient records in the current data set. This study was approved by the institutional review board of each hospital. Information was deidentified to protect patient confidentiality.

Variables Extracted

**Demographic and Medical Information**

Information was collected on the child’s age, gender, race, and zip code. The last variable was used to provide information on socioeconomic status. Through the US Census data, we identified median household income. In addition, the child’s length of hospitalization and diagnosis(ies) were gathered. Using *International Classification of Diseases, Ninth Revision*, diagnoses and codes, we identified patients with at least 1 psychiatric diagnosis (290–319), an autism spectrum disorder diagnosis (299), ADHD (314), and/or a cancer diagnosis (140–208).

**Medications**

Eight classes of medications (inclusive of 22 medications) were identified as potential sleep medications on the basis of previous studies and clinical experience. A list of these medications can be found in Table 1. Because there are no FDA-approved medications for pediatric insomnia, each of these medications could potentially have been prescribed for another reason. In this study, a medication was labeled as a “sleep medication” if it was (1) prescribed and administered on the specific target date as a once/daily dosing between 6:00 PM and 4:00 AM or (2) prescribed as needed but given only once that day between 6:00 PM and 4:00 AM.

**RESULTS**

**Sample Demographics**

The overall sample was 54.5% male, with a mean age of 7.0 years (SD: 6.3, range: 0–18 years inclusive). Children were 63% white, 26% black, 4% Hispanic, and 1.2% Asian. The median household income (based on zip code) was $44 402. In terms of diagnoses, 18.6% of patients had at least 1 psychiatric diagnosis, and 5.0% had a cancer diagnosis.

One-way analyses of variance and \( \chi^2 \) analyses were used to examine demographic differences between the 3 sites. Because of the large number of subjects and analyses conducted, a more conservative \( P \) value was set at .01 for all analyses. Significant differences were found between the sites for child’s age \((F_{2,9273} = 106.5; \ P < .001)\), median household income \((F_{2,9273} = 53.8; \ P < .001)\), child’s race \((\chi^2 = 764.9; \ P < .001)\), short (≤7 days) hospitalization versus longer hospitalization (>7 days) \((\chi^2 = 29.7; \ P < .001)\), and whether the child had a psychiatric diagnosis \((\chi^2 = 855.9; \ P < .001)\). A breakdown of the descriptive data for the demographic variables according to site is included in Table 2.

**Frequency of Prescribed Medications for Sleep**

The first aim of this study was to examine the frequency of sleep medications prescribed in hospitalized pediatric patients. Six percent of patients in this study were prescribed a medication from the list of potential sleep medications (Table 1). Of those prescribed sleep medications, 89% were prescribed 1 sleep medication, 10% were prescribed 2 sleep medications, and 1% were prescribed 3 sleep medications. As seen in Table 3, antihistamines were the most frequently prescribed medication (36.6%), followed by benzodiazepines (19.4%). Other common medications included antipsychotic agents (16.4%) and alpha agonists (10.4%). Nonbenzodiazepine hypnotic agents (eg, zolpidem, zaleplon) were the least frequently prescribed class of medications (2.2%).

**Differences in Frequency of Sleep Medications for Demographic and Medical Variables**

The second aim of the study was to examine whether there were any differences in prescribing patterns on the basis of demographic or medical variables. \( \chi^2 \) analyses were used to examine differences on demographic and

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Medication Class</th>
<th>Medication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha agonists</td>
<td>Clonidine</td>
</tr>
<tr>
<td>Antidepressants</td>
<td>Imipramine hydrochloride</td>
</tr>
<tr>
<td>SSRs</td>
<td>Citalopram</td>
</tr>
<tr>
<td>Antihistamines</td>
<td>Diphenhydramine</td>
</tr>
<tr>
<td>Antipsychotic agents</td>
<td>Olanzapine</td>
</tr>
<tr>
<td>Benzodiazepines</td>
<td>Alprazolam</td>
</tr>
<tr>
<td>Hypnotic agents</td>
<td>Zaleplon</td>
</tr>
<tr>
<td>Chloral hydrate</td>
<td>Zolpidem</td>
</tr>
</tbody>
</table>

---
medical variables for children who received a sleep medication and children who did not receive a sleep medication. As seen in Table 4, a significant difference was found for site ($\chi^2 = 102.3; P < .001$), with a lower percentage of patients at CHOP receiving sleep medications than the other 2 sites. No differences in the frequency of sleep medications was found on the basis of the child’s gender ($\chi^2 = 0.3, \text{not significant (NS)}$). To make the differences in age more meaningful, this variable was divided into 4 age groups (infant, toddler, child, and adolescent), and a significant difference was found for age ($\chi^2 = 232.8; P < .001$), suggesting that the frequency of sleep medications increases with increasing age. A significant difference in the frequency of sleep medications was also found for race ($\chi^2 = 19.7; P < .001$), with more white and black patients prescribed sleep medications than Hispanic and Asian patients. Hospital length ($\leq 7$ vs $> 7$ days) was also significantly different ($\chi^2 = 105.9; P < .001$), suggesting medications were more frequently prescribed during short-term hospitalizations.

In terms of diagnosis, a difference in whether the child was on a sleep medication was found depending on whether the child had a psychiatric diagnosis ($\chi^2 = 575.8; P < .001$), indicating that more children with a psychiatric diagnosis were given a sleep medication than those without a psychiatric diagnosis. Within the psychiatric diagnoses, children with autism spectrum disorders ($\chi^2 = 133.6; P < .001$) and ADHD ($\chi^2 = 200.6; P < .001$) were more likely than children without those disorders to have sleep medications prescribed. No difference was found for children with a cancer diagnosis ($\chi^2 = 0.5, \text{NS}$).

### Differences in Type of Sleep Medications for Medical Variables
A second set of $\chi^2$ analyses were used to examine differences in demographic and medical variables between the classes of sleep medications prescribed. A significant difference in the type of medications prescribed was found for the 3 hospitals ($\chi^2 = 68.8; P < .001$; see Table 3 for frequencies and percentages). Significant differences were also found for the service from which the child was discharged from ($\chi^2 = 198.7; P < .001$), whether the child had a psychiatric diagnosis ($\chi^2 = 575.8; P < .001$), indicating that more children with a psychiatric diagnosis were given a sleep medication than those without a psychiatric diagnosis. Within the psychiatric diagnoses, children with autism spectrum disorders ($\chi^2 = 133.6; P < .001$) and ADHD ($\chi^2 = 200.6; P < .001$) were more likely than children without those disorders to have sleep medications prescribed. No difference was found for children with a cancer diagnosis ($\chi^2 = 0.5, \text{NS}$).
Demographic and Medical Predictors of Prescribing Patterns

A logistic regression analysis was used to examine factors predicting the likelihood that a patient was given a sleep medication. Demographic and medical predictors were psychiatric diagnosis (presence or absence), child’s age, length of hospitalization (≤7 days or >7 days), race (white, black, Hispanic, Asian, other), and hospital service (general pediatrics, surgery, critical care, subspecialty, psychiatry). A test of the full model with all 5 predictors against a constant-only model was statistically significant ($\chi^2 = 579.9; P < .001$), indicating that the predictors, as a set, reliably distinguish between patients who were prescribed a sleep medication and those who were not. However, an examination of the Wald test revealed that race did not provide a significant predictor, thus the model was rerun without race, and remained statistically significant ($\chi^2 = 572.4; P < .001$). Table 6 shows regression coefficients, Wald statistics, odds ratios (ORs), and 95% confidence intervals (CIs) for ORs for each of the predictors in the final model. Patients were 2.7 times more likely to have a psychiatric diagnosis, twice as likely to have a short hospitalization, and 1.6 times more likely to be on the psychiatric service.

Secondary Chart Review at HCH

Because the proxy method may overestimate the frequency that medications are prescribed for sleep, a secondary chart review was conducted on the HCH charts, because this data collection approach allowed us to examine reasons why patients were given certain medications. Although 39 (4.8%) of 804 patients were identified as being on ≥1 medication for sleep, the secondary

### Table 4: Frequency and Percentage of Demographic Variables for Patients Prescribed Medications for Sleep

<table>
<thead>
<tr>
<th>Variable</th>
<th>Prescribed Sleep Medication</th>
<th>No Sleep Medication</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>%</td>
</tr>
<tr>
<td>Total sample</td>
<td>562</td>
<td>6.0</td>
</tr>
<tr>
<td>Site</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHOP</td>
<td>139</td>
<td>3.3</td>
</tr>
<tr>
<td>CHMC</td>
<td>364</td>
<td>8.3</td>
</tr>
<tr>
<td>HCH</td>
<td>59</td>
<td>7.5</td>
</tr>
<tr>
<td>Child's Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>244</td>
<td>5.7</td>
</tr>
<tr>
<td>Male</td>
<td>318</td>
<td>6.2</td>
</tr>
<tr>
<td>Developmental stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infant</td>
<td>30</td>
<td>1.1</td>
</tr>
<tr>
<td>Toddler</td>
<td>79</td>
<td>4.1</td>
</tr>
<tr>
<td>Child</td>
<td>198</td>
<td>8.5</td>
</tr>
<tr>
<td>Adolescent</td>
<td>255</td>
<td>10.2</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>398</td>
<td>6.7</td>
</tr>
<tr>
<td>Black</td>
<td>129</td>
<td>5.3</td>
</tr>
<tr>
<td>Hispanic</td>
<td>9</td>
<td>2.6</td>
</tr>
<tr>
<td>Asian</td>
<td>3</td>
<td>2.6</td>
</tr>
<tr>
<td>Other</td>
<td>20</td>
<td>3.9</td>
</tr>
<tr>
<td>Hospital length of stay</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤7 d</td>
<td>387</td>
<td>8.6</td>
</tr>
<tr>
<td>&gt;7 d</td>
<td>175</td>
<td>3.6</td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychiatric</td>
<td>319</td>
<td>18.2</td>
</tr>
<tr>
<td>No psychiatric</td>
<td>243</td>
<td>3.2</td>
</tr>
<tr>
<td>Autism</td>
<td>38</td>
<td>30.2</td>
</tr>
<tr>
<td>No autism</td>
<td>524</td>
<td>5.6</td>
</tr>
<tr>
<td>ADHD</td>
<td>106</td>
<td>20.2</td>
</tr>
<tr>
<td>No ADHD</td>
<td>456</td>
<td>5.1</td>
</tr>
<tr>
<td>Cancer</td>
<td>25</td>
<td>5.3</td>
</tr>
<tr>
<td>No Cancer</td>
<td>537</td>
<td>6.0</td>
</tr>
<tr>
<td>Service</td>
<td></td>
<td></td>
</tr>
<tr>
<td>General pediatrics</td>
<td>71</td>
<td>3.3</td>
</tr>
<tr>
<td>Surgery</td>
<td>69</td>
<td>3.3</td>
</tr>
<tr>
<td>Critical care</td>
<td>26</td>
<td>7.4</td>
</tr>
<tr>
<td>Subspecialty</td>
<td>149</td>
<td>4.0</td>
</tr>
<tr>
<td>Psychiatry</td>
<td>247</td>
<td>22.2</td>
</tr>
</tbody>
</table>

### Table 5: Frequency and Percentage of Demographic and Medical Variables for Classes of Medications Prescribed for Sleep

<table>
<thead>
<tr>
<th>Variable</th>
<th>Alpha Agonist</th>
<th>Antidepressant</th>
<th>SSRI</th>
<th>Antihistamine</th>
<th>Antipsychotic Agent</th>
<th>Benzodiazepine</th>
<th>Hypnotic Agent</th>
<th>Choral Hydrate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hospital service</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General pediatrics</td>
<td>9.3 (7)</td>
<td>5.3 (4)</td>
<td>0 (0)</td>
<td>32.0 (24)</td>
<td>9.3 (7)</td>
<td>32.0 (24)</td>
<td>1.3 (1)</td>
<td>10.7 (8)</td>
</tr>
<tr>
<td>Surgical</td>
<td>5.3 (4)</td>
<td>2.6 (2)</td>
<td>3.9 (3)</td>
<td>36.8 (28)</td>
<td>2.6 (2)</td>
<td>39.5 (30)</td>
<td>0 (0)</td>
<td>9.2 (7)</td>
</tr>
<tr>
<td>Critical care</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>21.4 (6)</td>
<td>10.7 (3)</td>
<td>50.0 (14)</td>
<td>10.7 (3)</td>
</tr>
<tr>
<td>Subspecialty</td>
<td>7.6 (13)</td>
<td>7.6 (13)</td>
<td>3.5 (6)</td>
<td>36.8 (63)</td>
<td>7.6 (13)</td>
<td>28.7 (49)</td>
<td>4.1 (7)</td>
<td>4.1 (7)</td>
</tr>
<tr>
<td>Psychiatry</td>
<td>14.7 (41)</td>
<td>6.5 (18)</td>
<td>8.3 (23)</td>
<td>39.2 (109)</td>
<td>28.1 (78)</td>
<td>1.8 (5)</td>
<td>1.1 (3)</td>
<td>0.3 (1)</td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychiatric</td>
<td>14.6 (53)</td>
<td>8.0 (29)</td>
<td>6.9 (25)</td>
<td>34.8 (126)</td>
<td>26.0 (94)</td>
<td>5.8 (21)</td>
<td>2.2 (8)</td>
<td>1.7 (6)</td>
</tr>
<tr>
<td>No psychiatric</td>
<td>4.5 (12)</td>
<td>3.0 (8)</td>
<td>2.6 (7)</td>
<td>39.1 (104)</td>
<td>3.4 (9)</td>
<td>38.0 (101)</td>
<td>2.3 (6)</td>
<td>7.1 (19)</td>
</tr>
<tr>
<td>Autism</td>
<td>27.3 (12)</td>
<td>11.3 (5)</td>
<td>9.1 (4)</td>
<td>20.5 (9)</td>
<td>25.0 (11)</td>
<td>4.5 (2)</td>
<td>2.3 (1)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>No autism</td>
<td>9.1 (53)</td>
<td>5.5 (32)</td>
<td>4.8 (28)</td>
<td>37.8 (221)</td>
<td>15.8 (92)</td>
<td>20.5 (120)</td>
<td>2.2 (13)</td>
<td>4.3 (25)</td>
</tr>
<tr>
<td>ADHD</td>
<td>23.4 (29)</td>
<td>4.0 (5)</td>
<td>6.5 (8)</td>
<td>33.1 (41)</td>
<td>29.8 (37)</td>
<td>2.4 (3)</td>
<td>0.8 (1)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>No ADHD</td>
<td>7.1 (36)</td>
<td>6.3 (32)</td>
<td>4.8 (24)</td>
<td>37.5 (189)</td>
<td>13.1 (66)</td>
<td>25.6 (119)</td>
<td>2.6 (13)</td>
<td>5.0 (25)</td>
</tr>
<tr>
<td>Cancer</td>
<td>3.0 (1)</td>
<td>0 (0)</td>
<td>3.0 (1)</td>
<td>48.5 (16)</td>
<td>9.1 (3)</td>
<td>21.2 (7)</td>
<td>15.2 (5)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>No cancer</td>
<td>10.8 (64)</td>
<td>6.2 (37)</td>
<td>5.2 (31)</td>
<td>36.0 (214)</td>
<td>16.8 (100)</td>
<td>19.3 (115)</td>
<td>1.5 (9)</td>
<td>4.2 (25)</td>
</tr>
</tbody>
</table>

Values presented are % (frequency).
chart review indicated that only about half of these patients \( (n = 18; 46\%) \) were actually taking these medications for sleep. The sleep medications used were clonidine \( (n = 4) \), diazepam \( (n = 2) \), diphenhydramine \( (n = 6) \), trazodone \( (n = 4) \), and zolpidem \( (n = 4) \).

**Secondary Analysis Without Patients With a Psychiatric Diagnosis**

Because some medications that are included in this study may be primarily used for psychiatric disorders, with the secondary benefit of improving sleep, the overall results of this study may be biased by the large percentage of patients with a psychiatric diagnosis. Thus, additional analyses were conducted across sites to examine the frequency of medications prescribed for sleep, excluding patients with a psychiatric diagnosis \( (n = 1754) \). Of these patients, 3.2% \( (n = 243) \) were prescribed a medication for sleep, with 91% of patients prescribed 1 medication and 9% prescribed \( \geq 2 \) medications for sleep. As seen in Table 7, antihistamines continued to be the most frequently prescribed medication \( (39.1\%) \); however, benzodiazepines were also prescribed commonly \( (38.0\%) \). Hypnotic agents continued to be the least frequently prescribed class of medications \( (2.3\%) \).

In terms of demographic and medical differences, differences continued to be found in the prescribing practices of the 3 sites, both in frequency \( (\chi^2 = 30.2; P < .001) \), and class of medications prescribed \( (\chi^2 = 30.4; P < .001) \), with the overall rate of prescriptions more similar between CHOP and CCHMC compared with HCH (see Table 7). Significant differences continued to be found for age \( (\chi^2 = 63.8; P < .001) \), with medication use increasing with chronological age, and length of hospitalization \( (\chi^2 = 27.9; P < .001) \), with medications prescribed more often for patients who have shorter hospitalizations. No differences were found for gender \( (\chi^2 = 0.03, \text{NS}) \) or race \( (\chi^2 = 12.6, \text{NS}) \) (see Table 8).

**DISCUSSION**

This study is the first to examine the use of sleep medications in hospitalized children. Overall, 6% of all hospitalized children receive some type of sleep medication, with 3% of all medically hospitalized children (not including children with a psychiatric diagnosis) being prescribed a sleep-related medication. Thus, ~1 of every 20 to 25 children are treated with a sleep medication while hospitalized.

In terms of type of medication, prescription antihistamines were the most commonly prescribed sleep medications, given to 37% of all children prescribed a sleep-related medication. The next most common medications are benzodiazepines \( (19\%) \), antipsychotic agents \( (16\%) \), and alpha agonists \( (10\%) \), with fewer children receiving antidepressants \( (6\%) \), selective serotonin reuptake inhibitors \( (5\%) \), chloral hydrate \( (4\%) \), and nonbenzodiazepine hypnotic agents \( (2\%) \). Differences existed in type of medication prescribed by both hospital and by service. Across all 3 children’s hospitals, antihistamines were most commonly prescribed; however, children at CCHMC were more likely to be given antipsychotic agents and alpha agonists, whereas benzodiazepines

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total Sample ( (n = 266) )</th>
<th>CHOP ( (n = 118) )</th>
<th>CCHMC ( (n = 95) )</th>
<th>HCH ( (n = 53) )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>%</td>
<td>Frequency</td>
<td>%</td>
</tr>
<tr>
<td>Alpha agonists</td>
<td>12</td>
<td>4.5</td>
<td>6</td>
<td>5.0</td>
</tr>
<tr>
<td>Antidepressants</td>
<td>8</td>
<td>3.0</td>
<td>3</td>
<td>2.5</td>
</tr>
<tr>
<td>SSRIs</td>
<td>7</td>
<td>2.6</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Antihistamines</td>
<td>104</td>
<td>39.1</td>
<td>53</td>
<td>44.9</td>
</tr>
<tr>
<td>Antipsychotic agents</td>
<td>9</td>
<td>3.4</td>
<td>7</td>
<td>5.9</td>
</tr>
<tr>
<td>Benzodiazepines</td>
<td>101</td>
<td>38.0</td>
<td>40</td>
<td>33.9</td>
</tr>
<tr>
<td>Hypnotic agents</td>
<td>6</td>
<td>2.3</td>
<td>3</td>
<td>2.5</td>
</tr>
<tr>
<td>Chloral hydrate</td>
<td>19</td>
<td>7.1</td>
<td>6</td>
<td>5.1</td>
</tr>
</tbody>
</table>
were almost 3 times as likely to be prescribed at CHOP and HCH.

As stated, differences in medication choice also differed by service and diagnosis. For example, critical care patients were much more likely to be given a benzodiazepine or hypnotic agent and less likely to be prescribed an antihistamine. Children with a psychiatric diagnosis were more likely to be prescribed an alpha agonist or antipsychotic agent, compared with higher use of benzodiazepines and chloral hydrate in children without a psychiatric diagnosis. More specifically, children with autism were much more likely to be prescribed an alpha agonist, an antidepressant, or an antipsychotic agent, whereas children with ADHD were more likely to be given an alpha agonist or an antipsychotic agent and much less likely to be prescribed a benzodiazepine. Furthermore, children with cancer were 15 times more likely to be prescribed a hypnotic agent and much less likely to be given an alpha agonist or chloral hydrate.

Thus, medication choice was quite varied across service and across hospital setting. Unfortunately there are no FDA-approved medications for pediatric insomnia, nor are there guidelines for medication choices,\textsuperscript{1,13} which is reflected in the wide range of medications actually given. In addition, there is vastly more information available regarding recommended guidelines for pharmacological management of adult insomnia, although less is known about appropriate treatment for hospitalized adults. For hospitalized adult patients, recent recommendations\textsuperscript{14} indicate using intermediate-acting benzodiazepines (eg, lorazepam, temazepam) as first-line agents, followed by nonbenzodiazepines (eg, zaleplon, zolpidem) as second-line agents because of their increased cost. Trazodone is also considered an appropriate choice. On the other hand, antihistamines, tricyclic antidepressants, barbiturates, and chloral hydrate are discouraged. These recommendations, albeit for adults, are quite counter to the use of specific agents prescribed to children, with antihistamines the most commonly used medications and nonbenzodiazepines the least used.

As indicated, a striking finding in this study was that children on a psychiatric service and/or with a psychiatric diagnosis are the ones most likely to receive such a medication. In this study, 22% of children on a psychiatric service received a sleep-related medication, and more specifically 30% of children with autism and 20% of children with ADHD received a sleep-related medication. This is in contrast to 7% of children in critical care and 3% of children in a general pediatric unit or surgical service. This finding is consistent with 1 study that found that adult psychiatric patients experienced worse sleep quality while hospitalized than surgical or medical patients.\textsuperscript{15}

There are some difficulties interpreting our findings because those medications indicated to be sleep-related by using our proxy method are also those medications used to treat psychiatric diagnoses. Thus, secondary analyses were conducted on just those children hospitalized with medical diagnoses, excluding all psychiatric diagnoses. Of these patients, 3.2% were prescribed a medication for sleep. Antihistamines continued to be the most frequently prescribed medication (39.1%); however, benzodiazepines were also commonly prescribed (38.0%). Hypnotic agents continued to be the least prescribed class of medications (2.3%).

Another interesting finding was that no differences for sleep-related medication use were found between general pediatric services and surgical patients. This finding is in contrast to a study of hospitalized adults that found that medical patients received more sedative medications and reported less sleep problems than did surgical patients (note that no psychiatric patients were assessed in this study).\textsuperscript{16}

There were also differences in medication use and medication chosen across hospital site. The differences in diagnosis accounted for much of the differences across hospital site, given that CCHMC has a large inpatient psychiatric unit, accounting for 25.3% of all of its hospitalized children, whereas CHOP and HCH do not have inpatient psychiatric units. However, even beyond this difference in each hospital’s population, there were ad-

\begin{table}[h]
\centering
\caption{Frequency and Percentage of Demographic Variables for Patients Prescribed Medications for Sleep for Patients Without a Psychiatric Diagnosis}
\begin{tabular}{lcc}
\hline
Variable & Prescribed Sleep Medication & No Sleep Medication \\
 & Frequency & % & Frequency & % \\
\hline
Total sample & 243 & 3.2 & 7443 & 96.8 \\
Site & & & & \\
CHOP & 106 & 2.7 & 3826 & 97.3 \\
CCHMC & 39 & 3.0 & 2927 & 97.0 \\
HCH & 48 & 6.5 & 690 & 93.5 \\
Child’s gender & & & & \\
Female & 108 & 3.1 & 3352 & 96.9 \\
Male & 135 & 3.2 & 4091 & 96.8 \\
Developmental stage & & & & \\
Infant & 30 & 1.1 & 2646 & 98.9 \\
Toddler & 68 & 3.8 & 1725 & 96.2 \\
Child & 65 & 3.8 & 1636 & 96.2 \\
Adolescent & 80 & 5.3 & 1427 & 94.7 \\
Race & & & & \\
White & 177 & 3.7 & 4609 & 96.3 \\
Black & 45 & 2.3 & 1936 & 97.7 \\
Hispanic & 8 & 2.5 & 314 & 97.5 \\
Asian & 3 & 2.8 & 105 & 97.2 \\
Other & 8 & 1.8 & 425 & 98.2 \\
Hospital length of stay & & & & \\
\leq7 d & 148 & 4.3 & 3260 & 95.7 \\
>7 d & 95 & 2.2 & 4181 & 97.8 \\
Service & & & & \\
General pediatrics & 51 & 2.5 & 1997 & 97.5 \\
Surgery & 62 & 3.2 & 1895 & 96.8 \\
Critical care & 23 & 7.5 & 284 & 92.5 \\
Subspecialty & 107 & 3.2 & 3264 & 96.8 \\
\hline
\end{tabular}
\end{table}
ditional differences in what specific medications are more or less likely to be given to children in each setting. For example, after excluding all children with a psychiatric diagnosis, hypnotic agents were more commonly used at HCH, whereas selective serotonin reuptake inhibitors and antipsychotic agents were prescribed at CCHMC, medications that were chosen much less frequently at CHOP and HCH.

As expected, older children and adolescents are more likely to be prescribed a medication for sleep, as well as children with shorter hospital stays. Age differences, however, are confounded by diagnosis, with children with a psychiatric diagnosis having an average age of 12.5 years compared with 5.8 years for all other hospitalized children. However, even when excluding children with a psychiatric diagnosis, older children continue to be more likely to be prescribed a sleep-related medication. Also, children with a shorter hospital stay are more likely to be given a medication. This result is consistent with the concept that although nonpharmacological interventions for pediatric sleep disturbances are effective, they may not be appropriate for hospitalized children with sleep problems because of the length of time needed before behavioral treatments can be effective.

Unfortunately, no studies to date have been conducted on the percentage of hospitalized children who experience sleep problems. Thus, it is difficult to assess whether hospitalized children and adolescents are getting medicated appropriately for sleep issues. An early adult study conducted in 1990, indicated that a large percentage of medical and surgical patients (range: 31%–88%) were prescribed a sedative hypnotic drug.17 A more recent study of elderly hospitalized patients found that 29% had a hypnotic-agent prescription initiated while hospitalized and an additional 31% continued a preadmission hypnotic-agent prescription while hospitalized.18 These findings are in strong contrast to the 3% to 6% found in this study.

There are several reasons why a lower rate of prescriptions for sleep problems were found in this study compared with previous chart reviews and community surveys. First, the community surveys reflected the prescribing practices for outpatient settings (both general pediatricians and pediatric psychiatrists), where sleep medications are typically prescribed for long-standing sleep problems, particularly for psychiatric patients, or acute needs (eg, travel, sleep disruption because of death in the family), which is not the case for hospitalized patients. Second, while children are at home, they are generally expected to sleep well, with only those sleep disruptions seen as “problematic” reported to pediatricians, resulting in the high prevalence of practitioners prescribing medications. However, most people expect that sleep during a hospitalization will be poor, because of pain, discomfort, being in a strange environment, or external noises (alarms, conversations in the hallway). Thus, sleep disruptions may not be seen as “problematic,” resulting in the lower rates of medications prescribed. Finally, the difference in medication profiles between the current study and previous studies is likely due to the rates of over-the-counter medications prescribed in outpatient settings. Because this chart review relied on pharmacy records, and over-the-counter medications for sleep are not recorded, this may also explain these differences in prescription practices.

There are several limitations to this study that may impact on the generalizability of these results. First, there are some concerns about the proxy method used. By using a proxy method endorsed by other studies on medication use for pediatric sleep problems,3 6% of all hospitalized children were classified as being prescribed a sleep medication. However, a more in-depth chart review at 1 of the study sites (HCH) found 39 (4.8%) of 804 children were administered a potential sleep medication from the list,19 but that a specific sleep indication could be subsequently definitively identified in only about half (n = 18; 46%) of those charts (alternative indications included pain, seizures, agitation, nausea, allergic reaction). Thus, the proxy method may somewhat overestimate the rate of sleep-medication prescriptions. However, the secondary chart review method may have underestimated the number of patients with specific sleep indications, because it is not always possible to assess exactly why a specific medication is given, and some of the medications may have been given for dual purposes (eg, a sedating antidepressant for depression and sleep [J.A.O., C.L. Rosen, MD, J.A.M., and L.H. Kirchner, MD, unpublished data, 2006] but only recorded for 1 reason, for example, depression).

A second limitation is that all 3 hospitals included in this study are academic institutions, thus they may not reflect prescription practices in community-based hospitals. A strength of this study was the inclusion of 3 hospitals, especially given that we found prescribing differences across the 3 hospitals that would not have been reflected if only 1 hospital was included. In addition to individual regional differences, the census of each hospital and the services provided had a clear impact on medication use. A study of other types of hospitals, especially community-based ones, would provide additional information on prescribing practices across a broader array of inpatient settings.

Overall, results from this study and other studies on prescription prevalence for sleep-related medications in children and adolescents indicate that prescribing medication for sleep in pediatric inpatients is somewhat common, although more so for children with psychiatric diagnoses and pediatric patients seen in outpatient settings. However, no FDA-approved medications exist, nor are there clinical consensus guidelines about type/specific medicines that should be used. Although our results
indicate variability in the types of medications prescribed, antihistamines and benzodiazepines are the most common ones currently used for hospitalized pediatric patients. In addition, the best predictors of a prescription for a sleep-related medication in a hospital setting are a psychiatric diagnosis, older age, and shorter hospital stay. The results of this study suggest the need for additional examination of and physician training focusing on sleep problems and potential interventions for hospitalized children. Furthermore, clinical trials and consensus on safe and efficacious medications for pediatric populations, including hospitalized children, are needed. Finally, more broad-based prevalence studies of pediatric insomnia in hospitalized children are necessary to more truly understand what is the need for pharmacological and behavioral treatments for these children.

ACKNOWLEDGMENTS
We thank Raymond Morris, Christine Gould, Juhee Lee, and Frank Baker for assistance with data collection.

REFERENCES

RANDOMIZED CLINICAL TRIALS
Pediatrics requires investigators to register their clinical trials in a public trials registry. The members of the International Committee of Medical Journal Editors (ICMJE) will consider most clinical trials for publication only if they have been registered (see N Engl J Med. 2004;351:1250–1251). Current information on requirements and appropriate registries is available at www.icmje.org/faq.pdf.
Compliance With Alarm Limits for Pulse Oximetry in Very Preterm Infants

Luisa Clucas, BMedSca, Lex W. Doyle, MD, FRCPa,b,c,d, Jennifer Dawson, RNc, Susan Donath, MAa,e, Peter G. Davis, MD, FRCPa,c

Departments of Obstetrics and Gynaecology and Paediatrics, University of Melbourne, Melbourne, Australia; Division of Newborn Services, Royal Women’s Hospital, Carlton, Australia; Clinical Epidemiology and Biostatistics Unit, Murdoch Children’s Research Institute, Royal Children’s Hospital, Melbourne, Australia

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

ABSTRACT

OBJECTIVE. The objective of this study was to determine the rate of compliance with hospital guidelines for alarm limits for pulse oximetry in preterm infants on oxygen therapy.

METHODS. All infants admitted to the nurseries at the Royal Women’s Hospital, Melbourne, Australia, with gestational age <32 weeks or birth weight <1500 g between August 2005 and February 2006 were eligible for inclusion. Data on the alarm limits set for infants on oxygen therapy were collected prospectively. The target saturation range recommended in written hospital guidelines was 88% to 92%, with alarm limits set at 85% and 94%.

RESULTS. Data were prospectively collected for 144 subjects with mean (SD) gestational age 29.3 (2.4) weeks and birth weight 1226 (354) g; 1073 alarm limits were collected when infants were on oxygen. The lower alarm limit was set correctly 91.1% of the time. In contrast, the upper alarm limit was set correctly only 23.3% of the time: 76.5% of the time it was too high, and 23.8% of the time it was set at 100%. Infants with an upper alarm limit set correctly on a particular day had a significantly lower birth weight, gestational age, postmenstrual age, and postnatal age than infants who had the upper alarm limit set too high. Use of assisted ventilation, higher inspired oxygen concentrations, and more frequent changes in inspired oxygen concentration were all associated with improved odds of having an appropriately set upper alarm limit.

CONCLUSIONS. This study suggests that current guidelines regarding the upper pulse oximeter alarm limit for infants receiving oxygen might be commonly exceeded, although compliance might be better for infants at higher risk of adverse outcomes. However, there might be less variation from guidelines for the lower alarm limit.
Oxygen, although vital to survival, can be toxic to living cells particularly in high concentrations, so much so that in preterm infants it is possible to have “too much of a good thing.” Oxygen toxicity in preterm infants has been associated with conditions such as bronchopulmonary dysplasia and retinopathy of prematurity (ROP). Unfortunately, preventing ROP is not as simple as restricting the amount of oxygen given to these infants: a balance between too much oxygen and too little must be found to reconcile the competing outcomes of ROP and bronchopulmonary dysplasia caused by too much oxygen, and neurologic impairment and death caused by too little. Monitoring the level of oxygen in the blood of these infants is, therefore, very important and is part of routine care.

Pulse oximetry is now the most widely used method for monitoring oxygen levels in these infants and is considered the standard noninvasive technique. There is, however, still no agreement on the optimum saturation range, and considerable variation exists between centers. At the Royal Women’s Hospital, Melbourne, the guidelines for the use of pulse oximetry in preterm infants receiving supplemental oxygen were changed in June 2005, partly in preparation for a masked randomized, controlled trial comparing 2 different oxygen saturation target ranges. There is little information about compliance with alarms limits for pulse oximetry for infants in oxygen, hence we sought to obtain this information for the most immature and tiny infants and to determine the clinical variables that affect compliance with guidelines. It was hypothesized that compliance with both upper and lower alarm limits would be good, and would be better in infants at highest risk.

**METHODS**

We conducted a prospective audit of the use of pulse oximetry in a tertiary neonatal center, the neonatal unit at the Royal Women’s Hospital. Infants were recruited from August 29, 2005, to February 28, 2006, and were eligible for participation if they had either a birth weight <1500 g or gestational age <32 completed weeks, and were admitted within the first day of life. Infants who left the unit within 24 hours of admission were excluded.

The guidelines for oxygen target ranges and alarm limits were changed on June 6, 2005 by the issuing of a memorandum to all staff, medical and nursing, involved in pulse oximetry. The new guidelines required target oxygen saturations between 88% to 92%, with alarm limits set at 85% and 94% for infants receiving supplemental oxygen. For infants not on oxygen, the upper alarm limit could be up to 100%. It was expected that all staff would follow the guidelines unless specifically ordered to do otherwise. There was no limit on how many changes were allowed to the inspired oxygen concentration in response to oxygen values outside the target range. Before June 2005, the target range was 90% to 95%, with lower alarm limits of 85% for infants <2 weeks old and 80% for older infants, and upper alarm limits of 96% for infants in oxygen.

Data for the current study were collected from August 29, 2005, until March 31, 2006. Data were collected at approximately the same time daily during weekdays for each infant during this period until the infant was discharged from the unit (and did not return within 24 hours), or until the end of the data collection period, whichever occurred first. Data included the length of time (hours) the infant was on an oximeter in the preceding 24-hour period, the targeted oxygen saturation range, and pulse oximeter alarm limits. If there was no specific saturation range ordered by the doctor in the patient’s treatment notes, it was assumed that the target range was 88% to 92% and that alarm limits should be set at 85% and 94% for all infants on supplemental oxygen, as per the hospital guidelines.

We recorded whether the infant was breathing air or supplemental oxygen. Only days when the infant was on supplemental oxygen are included in this study. The fractional inspired oxygen (FiO₂) (percent) or flow rate (mL/min) of oxygen at that time, the range of the FiO₂ or flow rate of oxygen (lowest and highest values) in the preceding 24-hour period, the number of times the amount of oxygen was changed, and whether the infant was on assisted ventilation at the time of data collection (any of continuous positive airway pressure, conventional ventilation, or high-frequency ventilation) were recorded. After all the data were collected, the oxygen requirements on any given day were categorized into approximate tertiles as low (FiO₂: 22%–23%), moderate (FiO₂: 24%–29%), or high (FiO₂ > 29%); FiO₂ values were derived for infants on low-flow oxygen before this categorization.

Ethics approval for this audit was obtained from the Royal Women’s Hospital Research and Ethics Committees.

**DATA ANALYSIS**

Data were analyzed by using SPSS 14.0 and Stata 9.1. Continuous data were expressed as mean (SD) or median (interquartile range [IQR]) if the data were skewed. Dichotomous data were presented as percentages.

All comparative analyses were adjusted for clustering, because of the repeated observations on the same subjects. For continuous variables (birth weight, gestational age, postmenstrual age, and postnatal age), linear regression with robust variance estimation (to adjust for clustering) was used to estimate the difference in mean values between days when the upper alarm limit was set at 94% and days when the limit was set at >94%. Multivariable logistic regression with robust variance estimation was then used to determine the independent effects of gestational age and postnatal age on the odds.
that the upper alarm limit would be set correctly; we did not want to test birth weight and postmenstrual age simultaneously in this analysis because they were too highly correlated with gestational age and postnatal age. χ² with adjustment for clustering was used to investigate whether categorical variables (gender, whether receiving assisted ventilation, FiO₂ tertile, and grouped number of changes in oxygen in the preceding 24-hour period) influenced the upper alarm-limit setting (94%, >94%) on each day. Multivariable logistic regression with robust variance estimation was used to estimate the independent effects of the last 3 variables on the upper alarm setting; results are presented as adjusted odds ratios (ORs) with 95% confidence intervals (CIs) on the basis of the robust variance estimate.

The sample size of this study was determined by the number of infants admitted to the unit within the study period who met the inclusion criteria, and by the number of days that those infants received oxygen therapy during their admission.

RESULTS

The study sample comprised 144 consecutive infants who met the inclusion criteria and who were born over the 6-month period between August 28, 2005, and February 28, 2006. This comprised all eligible infants admitted over this time. The subjects had a mean gestational age of 29.3 weeks (SD: 2.4) and mean birth weight of 1226 g (SD: 354). Seventy-seven (53.5%) infants were male. Infants were followed for a median of 31 days (IQR: 13.2–48.8). Of the 144 infants, 129 were followed until they were discharged from the unit or died, whereas 15 were still in the unit at the end of the data collection period. Some data for 11 of 144 infants were missed during the prospective data collection, having initially been excluded incorrectly on the basis of birth weight. Of 144 infants, 64 never received any oxygen. Of 80 who received oxygen, 55.6% (n = 44) were male with a mean gestational age of 28.4 weeks (SD: 2.4) and mean birth weight of 1123 g (SD: 383); they were in hospital for a median of 37 days (IQR: 17.0–61.5), during which time they had oximetry for a median of 32.5 days (IQR: 11.5–67.0) and were in oxygen for a median of 5 days (IQR: 2.0–34.5).

In total, 1073 lower and upper alarm limits were collected when infants were on oxygen. Of the 1073 lower alarm-limit values, 977 (91.1%) were set correctly at 85%, 6.3% were set lower than 85%, and 2.7% were set higher (Fig 1). Of the 1073 upper alarm-limit values, 250 (23.3%) were set correctly at 94%; 76.5% were set above 94%, and 23.8% were set at the maximum value of 100%. Only 2 of the upper alarm limits collected were set lower than 94% (0.2% of the values collected) (Fig 2). Of the 1073 pairs of lower and upper alarm limits collected when infants were on supplemental oxygen, there were 236 cases (22.0%) in which both the lower and upper alarm limits were set correctly. In 7.6% of cases both alarm limits were incorrect, and in 70.4% only 1 was correct.

The upper alarm limit was more likely to be set correctly on a particular day in infants of significantly lower birth weight, gestational age, postmenstrual age, and postnatal age (Table 1). Gestational age and postnatal age had independent effects; every week's decrease in gestational age increased the odds that the upper alarm limit would be correct by 36% (adjusted OR: 1.36; 95% CI: 1.20 to 1.54; P < .001), and for every week's decrease in gestational age increased the odds that the upper alarm limit would be correct increased by 11% (adjusted OR: 1.11; 95% CI: 1.06 to 1.16, P < .001).
Of the observations taken for infants requiring high levels of supplemental oxygen, the upper alarm limit was set correctly in 35.7% of cases, compared with 23.6% in the moderate oxygen group and 6.2% in the low oxygen group, a statistically significant difference between the groups (\(\chi^2: 20.8; P < .001\) adjusted for clustering).

The proportions of alarm limits set correctly was also related to the number of changes to an infant’s \(\text{FiO}_2\) on a given day (adjusted \(\chi^2: 21.4; P < .001\); 37.2% of infants with the highest number of changes (>12 changes) had correct upper alarm limit, compared with 23.1% of infants who had 6 to 12 changes, and 7.3% who had 0 to 5 changes.

A greater proportion of infants who were on assisted ventilation had their upper alarm limit set correctly compared with infants who were not on assisted ventilation (31.1% compared with 6.3%, adjusted \(\chi^2: 19.1; P < .001\)). Gender did not have an effect (adjusted \(\chi^2: 0.03; P = .87\)).

Being on assisted ventilation, receiving moderate or high levels of oxygen, and having a moderate (6–12) or high (>12) number of changes to infants’ \(\text{FiO}_2\) all independently increased the odds of having the upper alarm limit set correctly on a given day (Table 2).

**DISCUSSION**

This study found that the lower alarm limit for very preterm infants on supplemental oxygen was usually set correctly; however, the upper alarm limit was set too high the majority of the time. The percentage of times when both the upper and lower alarm limits were set correctly was only 22.0%, leaving at least 1 alarm limit set incorrectly 78.0% of the time. This was not what we expected to find and compares poorly with a study conducted by Laptook et al.\(^2\) in the United States in which, unannounced, intermittent surveys found pulse oximeter alarm limits to be set incorrectly only 26% (SD: 15%) of the time when lower and upper alarm limits were to be set at 80% and 96%, respectively, and 23% (SD: 16%) of the time when alarm limits were to be set at 80% and 94%. As in our study, they found that it was predominantly the upper alarm limit that was set incorrectly. They related this particularly to infants whose oxygen requirement varied from air to low levels of supplemental oxygen. This is similar to the results in our study; infants receiving lower concentrations of supplemental oxygen (\(\text{FiO}_2\): 22% or 23%) had the lowest proportion of upper alarm limits set correctly.

Although infants with a lower gestational age and postnatal age and those on ventilation or receiving higher levels of oxygen were more likely to have a correctly set upper alarm limit, as we hypothesized, the compliance was still poor. Moreover, although lower risk infants of increased maturity and postnatal age and those not on assisted ventilation may be considered less of a concern for staff than other infants, they are still at risk of developing ROP.

The results of our study suggest that improvement in
compliance with the guidelines for the upper pulse oximeter alarm limit for preterm infants requiring supplemental oxygen at our hospital is needed. The possibility that setting, and keeping, the upper alarm limit at 94% may not always be realistic may also need to be considered. However, the results of the study by Laptook et al. suggest that better compliance is possible. Because of a lack of other published studies on this topic of compliance with alarm limits, it is unclear whether our results or those of Laptook et al. are more typical.

Our study did not attempt to provide any evidence of a link between alarm-limit compliance and ROP rates. It was a descriptive study aiming to provide information about the use of oximeters in the unit. Our study was limited by the inability to download data from the oximeters; instead, we relied on snapshots of alarm settings rather than continuous recording of these variables. In a recent study, Hagadorn et al. collected oximetry data continuously for 72 hours each week for the first 4 weeks of life in 84 infants' gestational age from 14 different centers with varying upper and lower target ranges for oximetry. They reported that infants receiving modifiable oxygen had median \( \text{SpO}_2 \) values within the target ranges in 12 of 14 centers. Overall, infants spent 16% below, 48% within, and 36% above their NICU’s intended range. The studies of Laptook et al. and Hagadorn et al. and our own study indicate that there are problems with compliance with both target ranges and alarm settings. Clinicians need to be aware of these problems. Interpretation of the results of randomized, controlled trials targeting different saturation ranges will be problematic if these problems are not addressed; indeed, if there is not a clear separation of targeted ranges between groups in such trials, then it is possible that no benefit of either range may be found.

Data were collected once daily in our study, thus only the alarm limits at that time are known. The proportion of the time, or number of times per day, the alarm limits were correct or incorrect is unknown; only the number of times that the alarm limits were correct or incorrect at the time of data collection could be determined. Nursery staff were aware of the purpose of the study and the approximate time the observations were going to be made and had the opportunity to modify the alarm limits if they so desired; that the upper alarm limit was too high on the majority of the days that infants were in oxygen suggests that they did not avail themselves of this opportunity. We speculate that the total number of times alarm limits are incorrect in a day might be worse than we have described. Time spent within the target saturation range or alarm-limit range also could not be determined. Correct alarm limits may not mean the infant’s saturations are being kept in this range, whereas infants with incorrect alarm limits may in fact be saturating correctly. Additional research looking at whether the alarm limits accurately reflect the oxygen saturation range in which the infant is being nursed, and if better compliance with alarm limits improves oxygen targeting, would be of interest.

**CONCLUSIONS**

The current guidelines regarding the upper pulse oximeter alarm limit for very preterm infants in oxygen at the Royal Women’s Hospital are rarely followed. The majority of the time the upper alarm limits are being set too high, with a concerning proportion set at the maximum value of 100%. Compliance with the guidelines for the lower alarm limit is better, with little deviation from the guideline of 85%. Infants at higher risk of adverse outcomes from oxygen toxicity had an increased likelihood of having the alarm limit set correctly on a given day. Although this trend is somewhat reassuring, compliance rates for the upper alarm limits for these infants were still poor.

**REFERENCES**

11. *Intercooled Stata 9.1 for Windows* [computer program]. College Station, TX: Stata Corporation; 2005
A National Outbreak of *Ralstonia mannitolilytica* Associated With Use of a Contaminated Oxygen-Delivery Device Among Pediatric Patients

Michael A. Jhung, MD, MPH, Rebecca H. Sunenshine, MD, Judith Noble-Wang, PhD, Susan E. Coffin, MD, Keith St John, MS, Felicia M. Lewis, MD, Bette Jensen, MMSc, Alicia Peterson, MPH, John LiPuma, MD, Matthew J. Arduino, DrPH, Galit Holzmann-Pazgal, MD, Jane T. Atkins, MD, Arjun Srinivasan, MD

*Division of Healthcare Quality Promotion, Centers for Disease Control and Prevention, Atlanta, Georgia; †Department of Infectious Diseases, Children’s Hospital of Philadelphia, Philadelphia, Pennsylvania; ‡Acute Communicable Diseases and Emerging Infections, Philadelphia Department of Public Health, Philadelphia, Pennsylvania; †Deartment of Pediatrics and Communicable Diseases, University of Michigan Health System, Ann Arbor, Michigan; ‡Division of Pediatric Infectious Diseases, St Louis Children’s Hospital/Washington University School of Medicine, St Louis, Missouri; †Methodist Children’s Hospital of South Texas, San Antonio, Texas*

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

ABSTRACT

**OBJECTIVES.** In August 2005, the Centers for Disease Control and Prevention was notified of a *Ralstonia* species outbreak among pediatric patients receiving supplemental oxygen therapy with the Vapotherm 2000i (Vapotherm, Inc, Stevensville, MD). The Vapotherm 2000i is a reusable medical device that was used in >900 hospitals in the United States in 2005. *Ralstonia* are waterborne bacilli that have been implicated in hospital-acquired infections. We initiated an investigation to determine the source of the outbreak and implement infection control and prevention measures.

**PATIENTS AND METHODS.** We performed a case-control study at 1 hospital and conducted national case findings to obtain clinical and environmental samples for laboratory analysis. Case-patients had health care–acquired *Ralstonia* colonization or infection. Isolates were compared by using pulsed-field gel electrophoresis. We tested manufacturer-recommended disinfection protocols for the Vapotherm 2000i under simulated-use conditions.

**RESULTS.** Case-patients at the hospital (*n* = 5) were more likely to have received Vapotherm therapy than controls. Nationally, *Ralstonia mannitolilytica* was confirmed in 38 patients (aged 5 days to 7 years); 35 (92%) of the patients were exposed to the Vapotherm 2000i before recovery of the organism. Pulsed-field gel electrophoresis showed related *R mannitolilytica* strains from isolates sent from 18 hospitals in 12 states. A Vapotherm machine reprocessed with a protocol proposed by the manufacturer grew *Ralstonia* spp after 7 days of simulated use. In December 2005, Vapotherm recalled the 2000i.

**CONCLUSIONS.** Our findings suggest intrinsic contamination of Vapotherm devices with *Ralstonia* spp. New medical devices may provide therapy equivalent to current devices yet pose novel reprocessing challenges.
**RALSTONIA SPECIES ARE** Gram-negative bacilli that are commonly found in moist environments, such as water and soil, or on plants. They were implicated previously in health care-associated outbreaks, primarily because of contamination of patient care solutions.\(^1\)\(^-\)\(^3\) \textit{Ralstonia} are opportunistic human pathogens, particularly among immunocompromised patients,\(^4\)\(^-\)\(^6\) and are an infrequent cause of infection. From January through April 2005, \textit{Ralstonia} spp were recovered from respiratory cultures taken from 5 patients at hospital A, a pediatric hospital in Philadelphia, Pennsylvania. No \textit{Ralstonia} spp had been recovered from patients at this hospital in the previous 2 years. Before culture, 4 of 5 patients were treated with the Vapotherm 2000i (Vapotherm, Inc, Stevensville, MD), a device that delivers humidified, warmed oxygen via nasal cannula. This portable oxygen-delivery device uses a 0.01-\(\mu\)m reusable filter cartridge to separate its air and water compartments and was used widely by neonatal clinicians, with \(\approx\)5000 units used in \(>\)900 hospitals in the United States in 2005. A diagram of the Vapotherm device is shown in Fig 1.

The Vapotherm 2000i was originally cleared for marketing by the Food and Drug Administration (FDA) in August 2000 under a premarket-notification [510(k)] submission. The 510(k) process allows marketing of new medical devices on the basis of their comparability to legally marketed devices with the same intended use. This differs from the FDA’s premarket-approval (PMA) process for certain new or high-risk devices, which requires demonstration of a reasonable level of safety and effectiveness before approval.\(^7\) In 2000, almost 99% of new medical devices that were cleared for marketing were reviewed through the 510(k) process.\(^8\)

In August 2005, hospital A and the Pennsylvania and Philadelphia Departments of Health invited the Centers for Disease Control and Prevention (CDC) to assist in an investigation.

**PATIENTS AND METHODS**

**Hospital A Investigation**

We conducted a matched case-control study at hospital A to delineate risk factors for health care-associated \textit{Ralstonia} spp colonization or infection. A case-patient was defined as any patient from whom \(\geq 1\) \textit{Ralstonia} culture was recovered from a clinical specimen between December 1, 2004, and August 31, 2005. Patients in whom \textit{Ralstonia} was believed to be present before hospitalization were excluded.

Four controls were selected for each case. Cases and controls were individually matched on length of hospitalization. To ensure that patients with unrecognized \textit{Ralstonia} colonization were not selected as controls, we limited control selection to patients who had chart documentation of a respiratory culture that did not grow \textit{Ralstonia}. Information on potential risk factors for respiratory infection was abstracted from medical charts.

Assessment of the association between case status and categorical exposure variables was determined using the Cochran-Mantel-Haenszel statistic, with strata defined by match group. An adjusted odds ratio (OR) for each exposure was derived via the logit estimate; \(0.5\) was added to each cell because of small stratum-specific sample size,\(^9\) and tables with a 0 sum row or column were not included in the computation. Statistical analyses were performed by using SAS 9.1 (Statistical Analysis System, Cary, NC).

On the basis of observations of infection control practices at hospital A, environmental samples of potential sources of \textit{Ralstonia} spp were obtained and sent to the CDC. These included hospital potable water, ice from ice

---

**FIGURE 1**

Schematic of Vapotherm 2000i oxygen-delivery device. Water from the reservoir is pumped through a heater and then used to warm and humidify oxygen as it flows past the air compartment of the vapor-transfer (filter) cartridge. The air and water compartments are separated by a 0.01-\(\mu\)m hollow fiber membrane filter.
machines, sterile water, condensate from 2 mechanical ventilators, sterile respiratory solutions, used and unused Vapotherm filter cartridges, and surface swabs from 4 Vapotherm devices.

National Investigation
Because the Vapotherm 2000i is used in hospitals throughout the country, we conducted national case-finding via respiratory therapy, infectious disease, infection control, and public health communication networks to identify additional patients with Ralstonia spp colonization or infection. Institutions were asked to submit patient information and clinical and environmental isolates for identification and molecular testing. Probable cases were those in which Ralstonia was reported between January 1, 2005, and March 1, 2006, but no isolate was available for testing at the CDC. Confirmed cases had clinical isolates that were identified as Ralstonia spp in CDC laboratories. Differences between infected and colonized patients were measured by using the Mann-Whitney test statistic for continuous variables and the χ² test statistic for categorical variables.

Microbiology
Vapotherm filter cartridges were flushed with 45 mL of sterile water, phosphate-buffered saline, or Dey-Engley neutralizing broth (Becton Dickinson, Franklin Lakes, NJ). The resulting eluent was processed by membrane filtration through a 0.45-μm pore size, 47-mm mixed cellulose ester membrane filter (Millipore Corporation, Bellerica, MA) or by centrifugation at 3500 g for 15 minutes. Surface swabs of Vapotherm devices were obtained from air ports on the filter cartridge, air connections on the Vapotherm machine, and the machine port delivering humidified oxygen to the patient. Vapor samples were obtained from operating devices by connecting the oxygen-delivery tube to a sealed, sterile container and collecting the resulting condensate.

The CDC also cultured tap water used during the calibration step of new Vapotherm devices manufactured in Ireland. Environmental water and ice samples were collected with sodium thiosulfate to neutralize residual chlorine.

All samples were cultured on Trypticase soy agar with 5% sheep blood (Becton Dickinson), MacConkey II agar (Becton Dickinson), and R2A agar. Identification of isolate species was performed by using the Vitek 2 automated instrument (bioMérieux, Durham, NC) in combination with a series of standard biochemical tests.11

Molecular Typing
After digestion of chromosomal DNA with SpeI, molecular typing by pulsed-field gel electrophoresis (PFGE) was performed as described previously.12 PFGE patterns were compared by using Bionumerics 3.5 software (Applied Maths, Austin, TX), and isolates were considered related if Dice coefficients were >80% similar. For comparison, we obtained isolates of Ralstonia mannitolyltica, collected several years before the start of the current outbreak, from the Burkholderia cepacia Research Laboratory and Repository (University of Michigan, Ann Arbor, MI). PFGE analysis was also used to confirm species identification, demonstrating that Ralstonia pickettii isolates were genetically distinguishable from R. mannitolyltica isolates.

Test of Reprocessing Protocols
During the investigation, Vapotherm proposed 2 new disinfection protocols for its machines and cartridges. The CDC assessed the efficacy of each protocol on a machine that was known to be contaminated with Ralstonia. One protocol recommended circulating 200 ppm chlorine dioxide (ClO₂) in the device fluid path for 10 minutes with the filter cartridge in place. To test this protocol, vapor samples were obtained at 3 time points: before reprocessing, immediately after reprocessing, and after 7 days of continuous device operation outside of patient care areas.

The second protocol recommended circulating 1000 ppm ClO₂ in the device fluid path for 1 hour without a filter cartridge. This protocol was performed on a contaminated machine by Vapotherm personnel in Stevensville, MD; the machine was then shipped to the CDC. Vapor samples were obtained at 6 time points: before reprocessing, immediately after reprocessing, every 7 days for 21 days, and on day 30. During this period, the machine was run continuously in a CDC laboratory.

RESULTS
Hospital A Investigation
From January 1, 2004, to September 1, 2005, 5 case-patients were identified. Three patients had respiratory cultures that grew R. mannitolyltica; 2 patients had respiratory cultures that grew R. pickettii. Case-patient status was significantly associated with exposure to a Vapotherm device within 30 days before recovery of Ralstonia (OR: 18; 95% confidence interval: 2.2–140). No other exposures, including mechanical ventilation, were significantly associated with Ralstonia recovery (Table 1). Multiple Ralstonia spp were isolated from surface swabs taken from each of the Vapotherm devices that were tested (n = 4), whereas samples of hospital potable water, sterile water, ice, 2 mechanical ventilators, and sterile respiratory solutions did not grow Ralstonia. Ralstonia spp were also recovered from 10 (71%) of 14 used filter cartridges and 0 of 5 unused cartridges.

National Investigation
The CDC received reports of 45 patients from 20 hospitals in 16 states from whom Ralstonia spp were recovered. Nine (45%) of 20 hospitals reported multiple
cases of *Ralstonia* (mean: 4 patients; range: 2–7 patients), whereas the remaining 11 (55%) each reported a single case. Isolates from 38 patients (84%) were confirmed by the CDC as *R. mannitolilytica*; isolates were not available for 7 cases. All 38 confirmed cases were pediatric patients (age range: 5 days to 7 years), and 35 (92%) were exposed to Vapotherm before recovery of *Ralstonia*.

For patients exposed to Vapotherm, median exposure time was 9 days (range: 1–121 days), and median duration between last Vapotherm exposure and first positive culture was 4 days (range: 0–38 days). In 8 cases (21%), the reporting clinician indicated that *Ralstonia* had caused an infection, whereas the remaining cases were believed to represent colonization. Infected patients were significantly younger (*P* = .002) than colonized patients and were more likely to have fever >38°C (*P* = .031), develop leukocytosis according to the treating clinician (*P* < .001), display evidence of pneumonia on chest radiography (*P* = .034), and receive antibiotic therapy to treat *Ralstonia* (Table 2). One infection (3%) was reported by the treating physician to have complicated the course of an infant’s underlying lung disease, possibly contributing to the patient’s death. Confirmed cases were identified from January 2005 to January 2006, and 24 (63%) of 38 were reported after the field investigation at hospital A (Fig 2).

### Microbiology

The CDC received 134 clinical and environmental isolates from 22 hospitals in 13 states; 111 (83%) were identified as *R. mannitolilytica* (38 clinical, 73 environmental). Two hospitals reported recovery of *R. mannitolilytica* from unused Vapotherm cartridges. One facility recovered *R. mannitolilytica* from 3 of 3 unused cartridges from a single lot; the other facility recovered *R. mannitolilytica* from 3 of 10 unused cartridges from 3 different lots. The CDC tested 26 new cartridges from 13 separate lots (including 2 lots from which contaminated new cartridges were reported) and failed to isolate any organisms.

Microbiologic analysis of samples taken from the Ireland facility where Vapotherm machines were calibrated yielded *Sphingomonas paucimobilus* from cultures of tap water and *B. cepacia* from swabs of the tap from which the water was drawn.

### Molecular Typing

Dice coefficients generated from analysis of the PFGE data for 111 isolates of *R. mannitolilytica* ranged from 100% to 59% relatedness. Clinical and environmental isolates from 18 (82%) of 22 hospitals in 12 noncontiguous states (Fig 3) were related, including isolates from 31 confirmed cases (82%). None of the 9 reference *R. mannitolilytica* isolates obtained before the outbreak began

---

**TABLE 1** Adjusted ORs for *R. mannitolilytica* Colonization in Hospital A, Pennsylvania, December 1, 2004, to August 31, 2005

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Cases (N = 5), n (%)</th>
<th>Controls (N = 20), n (%)</th>
<th>OR (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vapotherm use within 30 d</td>
<td>4 (80)</td>
<td>1 (5)</td>
<td>18 (2.2–140)</td>
</tr>
<tr>
<td>Ventilator use within 7 d</td>
<td>4 (80)</td>
<td>18 (90)</td>
<td>0.3 (0.01–8.2)</td>
</tr>
<tr>
<td>Nasogastric feeding</td>
<td>3 (60)</td>
<td>10 (50)</td>
<td>2.0 (0.1–31)</td>
</tr>
<tr>
<td>Inhaled medication</td>
<td>3 (60)</td>
<td>13 (65)</td>
<td>0.8 (0.1–3.9)</td>
</tr>
<tr>
<td>Cardiac ICU stay</td>
<td>3 (60)</td>
<td>3 (15)</td>
<td>5.1 (0.71–38)</td>
</tr>
<tr>
<td>Diagnosed cystic fibrosis</td>
<td>1 (20)</td>
<td>0 (0)</td>
<td>4.7 (0.75–29)</td>
</tr>
</tbody>
</table>

* Logit estimate for Cochran-Mantel-Haenszel test statistic.

---

**TABLE 2** Differences Between Infected and Colonized Confirmed Cases for Vapotherm-Associated *R. mannitolilytica* Outbreak, 2005–2006

<table>
<thead>
<tr>
<th></th>
<th>Infected (N = 8), n (%)</th>
<th>Colonized (N = 30), n (%)</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median age, d</td>
<td>14</td>
<td>56</td>
<td>.002</td>
</tr>
<tr>
<td>Male</td>
<td>3 (38)</td>
<td>17 (59)</td>
<td>.289</td>
</tr>
<tr>
<td>Vapotherm exposure before culture</td>
<td>7 (88)</td>
<td>28 (93)</td>
<td>.587</td>
</tr>
<tr>
<td>Premature birth</td>
<td>6 (75)</td>
<td>23 (77)</td>
<td>.922</td>
</tr>
<tr>
<td>Fever &gt;38°C</td>
<td>4 (50)</td>
<td>3 (13)</td>
<td>.031</td>
</tr>
<tr>
<td>Leukocytosis</td>
<td>6 (75)</td>
<td>2 (9)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Pneumonia by chest radiography</td>
<td>5 (63)</td>
<td>5 (22)</td>
<td>.034</td>
</tr>
<tr>
<td>Antibiotic therapy for <em>Ralstonia</em></td>
<td>8 (100)</td>
<td>8 (35)</td>
<td>.001</td>
</tr>
</tbody>
</table>

*a* Mann-Whitney test used for age variable; *χ²* test used for all other variables.

*b* One missing value.

*c* Seven missing values.
was related to these isolates. Hospitals from 10 additional states reported recovery of *Ralstonia* spp but either did not submit an isolate for confirmation or submitted isolates that were not related.

**Testing of New Reprocessing Protocols**

No organisms were recovered from the Vapotherm device tested immediately after disinfection with 200 ppm ClO₂. However, *R. mannitolilytica* grew in the device 7 days after reprocessing. After disinfection with 1000 ppm ClO₂, fewer than 10 cfu/mL of *Methylobacterium* species were recovered from samples taken immediately after reprocessing and on day 3. No organisms were recovered from cultures performed 7, 14, 21, or 30 days postreprocessing.

**DISCUSSION**

Our investigation demonstrates that contamination of a respiratory gas humidification device resulted in the recovery of *R. mannitolilytica* from patients in hospitals in 22 states. The organism could not be eradicated by using the manufacturer’s recommended disinfection protocol. Although generally of low virulence, *Ralstonia* spp can cause potentially serious health care-associated infections.³,¹³,¹⁴ This investigation, which epidemiologically and microbiologically linked transmission of *Ralstonia* to the Vapotherm 2000i humidification device, represents the largest reported outbreak associated with *R. mannitolilytica* known to the authors. Genetically related strains in disparate geographic locations suggest intrinsic contamination of the Vapotherm 2000i as the cause of the outbreak (Fig 3), although the source of this contamination has not been identified. Because water is used to manufacture or calibrate both the Vapotherm device and its filter cartridge, contamination during the production of 1 these components represents the most likely explanation for this outbreak.

Intermittent cartridge contamination is 1 possibility. Although the CDC did not recover organisms from the unopened cartridges it tested, 2 hospitals reported recovery of *Ralstonia* spp from unopened filter cartridges. However, cartridges are manufactured by using deionized water subject to reverse osmosis and subsequently undergo high temperature drying and annealing (150°F for 4 hours, followed by 250°F for 40 minutes), which makes introduction and survival of organisms unlikely.

Another explanation for the outbreak is the introduction of *R. mannitolilytica* into Vapotherm components during device calibration. Before the investigation, tap water was used to calibrate devices during manufacturing, and they were packaged for distribution before drying completely. Organisms in residual water in the machines could have formed biofilms in device tubing during shipping and storage,¹⁵ rendering them less susceptible to disinfection.¹⁶ *Ralstonia* spp are known to exhibit biofilm formation in plastic water piping.¹⁷ Once present in the devices, *Ralstonia* may have persisted because of a combination of factors, including warm and moist operating conditions that would promote bacterial growth, failure of end users to reprocess the devices on a consistent basis, and inability of the manufacturer’s original disinfection process to eradicate biofilms.

The results of this investigation are subject to several limitations. First, the source of Vapotherm device contamination was not identified. The CDC was unable to recover organisms from unopened cartridges, and *R. mannitolilytica* were not recovered from water samples taken from the manufacturing plant. However, other waterborne microorganisms, including *S. paucimobilis* and *B. cepacia*, were found in the water source used for cali-
bration. Water samples sent to the CDC were collected several months after the proposed inoculation had occurred, long enough for the microbial flora of the water to have changed, obscuring evidence of *Ralstonia* contamination at an earlier time point.

A second limitation stems from the limited information describing the genetic diversity of *R. mannitolilytica* in the United States. If *R. mannitolilytica* has only marginal diversity, then the finding of genetically related strains in several states may not imply intrinsic contamination of the Vapotherm device. Because *Ralstonia* spp are ubiquitous environmental organisms, the outbreak could have resulted from concomitant but unconnected contamination with apparently related strains. However, our investigation helps show that such genetic uniformity is unlikely. Twenty-two percent of tested *R. mannitolilytica* isolates showed PFGE patterns that were <80% related to the predominant strain. Furthermore, comparison isolates of *R. mannitolilytica* obtained before the start of the outbreak were not related to isolates collected during the investigation.

The Vapotherm 2000i was cleared for marketing through the premarket-notification 510(k) process with the indication, “To add moisture to, and to warm breathing gases for administration to patients.” Unlike PMA submissions, which require data that demonstrate the safety and efficacy of devices, manufacturers of devices submitted for 510(k) review are generally not required to submit design and manufacturing test data to the FDA. Instead, a 510(k) submission relies on demonstration of “substantial equivalence” to a “predicate device” that has already been cleared for marketing. A new device is considered substantially equivalent if it has the same intended use and technologic characteristics as the predicate device and does not raise new questions of safety and effectiveness. Device manufacturers seeking to market through the 510(k) process can claim reprocessing efficacy if their products are “virtually identical from an infection control perspective to a predicate device for which disinfection has been validated.” The FDA granted the Vapotherm 2000i clearance on the basis of a comparison to other marketed respiratory gas

![Dendrogram of *R. mannitolilytica* isolates: multiple states, 2005–2006. Hospitals with related isolates are shown with unrelated strains for comparison. Multiple states contained hospitals that had both related and unrelated strains of *R. mannitolilytica*. Isolates were obtained from the University of Michigan Burkholderia cepacia Research Laboratory.](image-url)
Contamination of either the Vapotherm device and/or vapor transfer cartridge, a component of machine interior during initial calibration failure to remove organisms during routine decontamination.

However, reprocessing instructions for reusable components of the Vapotherm device recommended use of a low-level disinfectant between patients. Furthermore, because the Vapotherm 2000i used a reusable 0.01-μm filter cartridge as a "biological barrier" between air and water compartments, tap water was permitted in the device. In its 510(k) Indication for Use Statement, the company stated, "under normal working conditions, there was essentially no risk of bacterial contamination of nasal airflow."23

On the basis of the results of this investigation, the FDA issued a Preliminary Public Health Notification on December 20, 2005, advising health care providers to use alternate devices. On December 22, 2005, Vapotherm recalled the 2000i device. This action was classified by the FDA as a Class I recall, which is reserved for products that have a reasonable chance of causing serious health problems or death.

In response to the outbreak and investigation, Vapotherm introduced changes to their manufacturing process and disinfection and use parameters to address potential sources of device contamination (Table 3). In January 2007, the FDA determined that these modifications satisfied requirements for a 510(k) submission, and the Vapotherm device was reintroduced for use. The CDC and FDA recommend that clinicians intending to use the Vapotherm device follow the latest instructions for the device and its components, which can be obtained by contacting Vapotherm.

This investigation emphasizes the importance of careful attention to infection control principles during the development, regulatory review, and use of medical devices. New devices may perform equivalently to currently marketed devices, but may also present novel reprocessing challenges when used in clinical settings. In recognition of this, the FDA has identified device sterilization as 1 of 3 key areas for improvement for 510(k) submissions.24 All medical devices have the potential to become contaminated during use, and thereby carry the potential for transmission of infectious diseases. To protect patients from device-associated infections, manufacturers should ensure that reusable devices can be consistently and effectively reprocessed between users. Reusable medical device users should also be aware that review of validation data for reprocessing methods is generally not included as part of the current 510(k) premarket-notification process. Regardless of how devices are cleared for marketing, end users are encouraged to review infection control and reprocessing guidelines under actual-use conditions to help ensure patient safety when using medical devices.

**ACKNOWLEDGMENTS**

Our outbreak investigation group included Corey Robertson, MD (New Jersey Department of Health and Senior Services); Corinne Walentik, MD, Morey Gardner, MD, Brenda Hinson, RN, Amie Keck, RT, and Fran Hixson, RN (St Mary's Health Center, St Louis, MO); James Ripka, RRT, Claudia Castellon, RN, and Lawrence Ross, MD (Children's Hospital Los Angeles, Keck School of Medicine, Los Angeles, CA); Daniel New, MD, Lori Patterson, MD, Sheila Ware, RRT-NPS, and Caroline Graber, RN (East Tennessee Children's Hospital, Knoxville, TN); Susan A. Dolan, RN, John F. James, PhD, and Trent Lucas, RRT (Children's Hospital, Denver, CO); Abbott Laptook, MD (Women and Infants' Hospital of Rhode Island, Providence, RI); Barbara Stein, RN (Children's Hospital of The King's Daughters, Norfolk, VA); Jane Harper, MS, RN, and Kathleen Harriman, PhD (Minnesota Department of Health, St Paul, MN); Michelle Hulse, MD, and Jane Harper, MS (Children's Hospitals and Clinics of Minnesota, Minneapolis, MN); Beth

<table>
<thead>
<tr>
<th>Potential Sources of Intrinsic Contamination of Vapotherm System and Actions Taken by the Manufacturer in Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potential Cause of Contamination</td>
</tr>
<tr>
<td>Contamination of machine interior during initial calibration with unfiltered water</td>
</tr>
<tr>
<td>Contamination of vapor transfer cartridge, a component of the device, during manufacture</td>
</tr>
<tr>
<td>Contamination of either the Vapotherm device and/or vapor transfer cartridge during patient use</td>
</tr>
<tr>
<td>Failure to remove organisms during routine decontamination</td>
</tr>
<tr>
<td>Correcive Actions</td>
</tr>
<tr>
<td>All devices in distribution were recalled to the manufacturer's facility and disinfected with 1000 ppm ClO₂.</td>
</tr>
<tr>
<td>Manufacture of new devices requires use of filtered water, followed by 1000 ppm ClO₂ disinfection and a dedicated drying step.</td>
</tr>
<tr>
<td>Manufacture of all new vapor transfer cartridges now includes ethylene oxide sterilization.</td>
</tr>
<tr>
<td>The company has developed a new system for retaining sterile water and now recommends that only sterile water be used for humidification.</td>
</tr>
<tr>
<td>The vapor transfer cartridge, previously a multi-use item, is now recommended for use by a single patient and should be replaced after 30 d of use.</td>
</tr>
<tr>
<td>The company continues to recommend disinfection of the Vapotherm unit between patients or after every 30 d of use, and now offers a peracetic acid formulation in addition to the quaternary ammonium disinfectant previously packaged with the device.</td>
</tr>
</tbody>
</table>

* Previously, the device used a refillable reservoir bag that created an open water circuit and allowed the use of tap water for humidification.
Ann Ayala, MS (Christus Santa Rosa Children’s Hospital, San Antonio, TX); Cindy Shifflett, BSN, Susan Latham, MT (ASCP), and Kim Cantrell, RRT-NPS (Spartanburg Regional Health Services District, Inc, Spartanburg, SC); Lynn Semproch, MT (ASCP) (Children’s Hospital of Michigan, Detroit, MI); Joyce Hayes-Berkowitz, RN, Mary Fairchild, RRT-NPS, and Kathy Acton, MT-ASP (Cardinal Glennon Children’s Hospital, St Louis, MO); Mary Rotar, RN (Children’s Hospital of Wisconsin, Milwaukee, WI); Craig H Gilliam, BSMT, Nicholas DePalo, BSRRT, Michele Honeycutt, RN, Toni Beavers-May, BSMT (ASCP), Toben Bennett, MSE, Stefan Juretschko, PhD, Gordon Schutze, MD (Arkansas Children’s Hospital, Little Rock, AR); Linda Campagna, RN, Drucy Borowitz, MD, Diane M. Dryja, MT (ASCP) (Women and Children’s Hospital of Buffalo, Buffalo, NY); Janet Eagan, RN (Memorial SloanKettering Cancer Center, New York, NY); Patti Kieffer, RN, Karla Howell, RN, Marla Ovart, CRT, and Kathy Long, CRT (Children’s Hospital of St Louis, St Louis, MO); Jolianne Stone Tocco, MPH (Oklahoma State Department of Health); Jeneene Kitz, BSN (OU Medical Center, Oklahoma City, Oklahoma); Barbara Quattlebaum, RN, and Patricia Hudak, MT (Methodist Children’s Hospital of South Texas, San Antonio, TX); Tam Ly, MD, Zacharia Cherian, MD, Joseph Lynott, MS, Nancy Donegan, MPH (Washington Hospital Center, Washington, DC).

REFERENCES


Temporal Relationships Between Colds, Upper Respiratory Viruses Detected by Polymerase Chain Reaction, and Otitis Media in Young Children Followed Through a Typical Cold Season

Birgit Winther, MD, PhD, Cuneyt M. Alper, MD, Ellen M. Mandel, MD, William J. Doyle, PhD, J. Owen Hendley, MD

Departments of Otolaryngology and Pediatrics, University of Virginia Health System, Charlottesville, Virginia; Department of Otolaryngology, Children’s Hospital of Pittsburgh and University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania

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

ABSTRACT

INTRODUCTION. Otitis media is a frequent complication of a viral upper respiratory tract infection, and the reported co-incidence of those diseases increases with assay sensitivity and sampling density. We determined the incidence of otitis-media complications in young children when referenced to cold-like illnesses and to concurrent virus recovery from the nasopharynx.

METHODS. A total of 60 children from 24 families were followed from October 2003 through April 30, 2004, by daily parental recording of illness signs, weekly pneumatic otoscopic examinations, and periodic polymerase chain reaction assay of collected nasal fluids for common viruses.

RESULTS. One hundred ninety-nine cold-like illnesses were observed, but a sample for virus assay was not collected concurrent with 71 episodes. Of the remainder, 73% of cold-like illnesses were temporally related to recovery of 1 or a combination of the assayed viruses, with rhinovirus predominating. For non–cold-like illness periods, 54 (18%) of 297 assays were positive for virus, and the virus frequency distribution was similar to that for cold-like illnesses. There were 93 diagnosed otitis-media episodes; 65 (70%) of these occurred during a cold-like illness. For the 79 otitis-media episodes with available nasal samples, 61 (77%) were associated with a positive virus result. In this population, the otitis-media complication rate for a cold-like illness was 33%.

CONCLUSIONS. A cold-like illness was not a prerequisite for polymerase chain reaction detection of viruses in the nose and nasopharynx of young children. Viral detection by polymerase chain reaction in the absence of a cold-like illness is associated with complications in some subjects. Otitis media is a complication of viral infection both with and without concurrent cold-like illnesses, thus downwardly biasing coincidence estimates that use cold-based illnesses as the denominator.
I t is commonly accepted that viral upper respiratory tract infection (vURI) is causally related to the development of new episodes of otitis media (OM). However, disagreements exist as to the relative importance of the different viral species in promoting the complication, true complication rate, and relative percentage of all OM episodes that are complications of a vURI. Resolution of these issues is important given that many believe OM to be preventable by using strategies to reduce vURI burden and/or to decouple OM as a presentation.

Typically, the presence of a vURI is recognized based on a set of self-appreciated and/or secondarily assigned symptoms and signs that define a cold-like illness (CLI). Because vURIs without symptom/sign expression do not present clinically, co-incidence assessments for the different vURI complications are referenced usually to CLIs. For OM, past estimates suggest that 20% to 40% of CLIs in children are complicated by OM and that 50% of all new OM episodes are attributable to a CLI. Recent studies show that these estimates are increased by increasing the density of otologic assessments (ie, capturing more OM events). Moreover, experimental infection of adults with different upper respiratory viruses shows that one half of those who are judged to be infected by shedding and/or seroconversion do not express signs/symptoms of a CLI. In 1 analysis of those data, the frequency of otologic complications in non-ill subjects remained significant, and other studies showed that preexisting antibodies and antiviral treatment differentially affected the CLI and otologic complications. Also, the presence of viruses in the nose and nasopharynx without a concurrent CLI was documented in control populations of infants, children, and adults, although the significance of this observation to presentations of complications such as OM was not assessed.

The purpose of this study was to determine whether subclinical vURIs are complicated by OM in a group of young children followed by daily diary for illness signs, weekly pneumatic otoscopic examinations by study personnel, and periodic collection of nasal secretions for virus assay by polymerase chain reaction (PCR). The null hypotheses tested are that the OM complication rate for a vURI estimated from CLI is an accurate estimate of the true complication rate for vURIs and that the ratios of OM coincidence to total OM episodes are similar when coincidence is defined for CLIs and vURIs. Acceptance or rejection of these hypotheses is important to the design of strategies for preventing OM on the basis of interfering with virus acquisition, virus shedding, or illness presentation during a vURI by providing estimates of the maximum efficacies and efficiencies of different proposed interventions.

MATERIALS AND METHODS
The data for this report were abstracted from those available for the first year of our ongoing, 5-year study entitled “Role of Virus and Genetic Susceptibility in Otitis Media.” In the study, families with 2 children aged 1 to 5 years are recruited by newspaper advertisement for participation in each year. Exclusion criteria include the presence in either child of a serious medical condition, a medical condition that predisposes to persistent OM, a nonintact or structurally abnormal tympanic membrane, a preexisting sensorineural hearing loss, or an inability to cooperate sufficiently with the examination and test procedures. After affirmation of willingness to participate and acquisition of written informed consent, families are entered into the study in October and followed through April of the respective study year. Families are reimbursed $100 per month for their participation. The study protocol was approved by the institutional review boards at the University of Pittsburgh and the University of Virginia.

The data for this report include general demographic information on the family obtained at entry, and the results for daily diaries maintained by the mother rating the presence/absence of 6 signs (ie, runny nose, nasal congestion, sore throat, cough, fever, and irritability) and recording the presence/absence of a cold or flu on the basis of their usual diagnostic criteria in each of their children; weekly assessments of middle-ear status in the children by pneumatic otoscopy done at either an “in-home visit” (Pittsburgh) or at the study clinics (Charlottesville) by study personnel; and periodic collection of nasal secretions from the children for virus identification by PCR performed at either an “in-home visit” or at the study clinics by study personnel. The purposes of our analyses were to determine the OM complication rate for a CLI and for times of recovery of an upper respiratory virus, and conversely, the frequencies of all OM episodes that are complications of a CLI or of recovery of an upper respiratory virus.

To define a CLI, we used a previously developed algorithm that reduces the between-rater biases in cold assignment. Specifically, the different mothers used a wide variety of sign constellations for assigning a cold or flu, and the algorithm standardizes the definition of a CLI across families. The algorithm first assigns an illness day if the maternal report for that day included ≥2 of runny nose, nasal congestion, and cough, or if the parental report included runny nose or nasal congestion given that the previous day was assigned by algorithm to an illness day. Then, illness/nonillness days were linked as child-specific strings for sequential observations beginning in October and extending to April 30, with string elements coded as 0 (nonillness day) or 1 (illness day). For the few times where the string sequences were broken by missing values, a default value of 0 was entered unless the bordering observations indicated illness.

1070 WINTHER et al
days wherein a value of 1 was entered. All zero-valued elements were deleted from each string to yield variable-length clustered sequences of 1s, which if at least 3 days in length, was considered to be a CLI, with onset defined as the first observation of an illness day and termination as the first nonillness day in a sequence of at least 5 consecutive nonillness days.

Data for bilateral pneumatic otoscopy performed at each examination were classified dichotomously for each ear as OM absent/present on the basis of the otoscopyist’s ratings of the tympanic membrane with respect to visibility, condition, position, appearance, color, vascularity, light reflex, and mobility. Data were coded for the left and right ears as OM absent/present (0 = absent, 1 = present) without consideration to the concurrent expression of otologic signs/symptoms. For each child, observation days for the right and left ears were linked as temporal strings, with interobservation values assigned the value of the initial observation for the interval. An OM episode was defined as a sequential series of 1s bounded by 0s for a given ear.

We proposed to collect periodic nasal secretion samples at monthly intervals and supplemental samples during periods of “parent-identified” cold episodes and/or otoscopically diagnosed OM. This proved not to be possible because some children resisted sample collection in the absence of free secretions (usually non-CLI periods), the failure in some instances of parents to promptly notify study personnel of a cold diagnosis in their child, the unavailability of families for a study visit during a parent-identified cold (eg, holiday, vacation), and the larger than expected number of parent-identified colds, which required reducing the number of samplings during non-CLI periods for budgetary reasons. A total of 513 samples were collected and assayed for viruses.

The technique for collecting secretion from nose and nasopharynx was described previously. Briefly, a Yankauer suction device (Conmed, Utica, NY) connected to a Lukens specimen trap (Sherwood Medical, St Louis, MO) was positioned sequentially in each vestibule to aspirate nasal secretions with application of negative pressure (≈90 mm Hg) from a suction pump. If visible secretion did not reach the tubing, when tolerated by the subject, 1 to 2 mL of nonbacteriostatic saline was instilled into the nostril with aspiration from the contralateral nostril. Visible secretion in the tubing was washed into the trap with saline.

All collected specimens were frozen at −70°C and shipped in batches to our virology laboratory at the University of Virginia for PCR assay to detect adenovirus, coronavirus, influenza virus, parainfluenza virus, rhinovirus (enterovirus), and respiratory syncytial virus (RSV). These viruses were sought by using a protocol adapted from the commercially available Hexaplex procedure (Prodesse, Inc, Waukesha, WI).27 Viral nucleic acid was extracted from 280 μL of aspirate/wash sample using the QIAmp Viral RNA Mini-Kit (Qiagen, Valencia, CA). Complementary DNA (cDNA) was synthesized by using random hexamers and reverse transcription. Viral nucleic acid was purified and reverse transcribed to cDNA when original samples were thawed for the first time. The quantity of cDNA was dictated by the planned number of different PCR amplifications, because testing of refrozen original samples or purified RNA may give false-negative results.25 Purification of adenoviral DNA by this technique (including the superfluous reverse transcription) was equivalent to extraction by a method for DNA (QIAmp DNA Blood Mini-Kit; Qiagen), which was used.28

PCR was performed with master mix and amplification parameters specific for each virus. Unincorporated primers and deoxynucleoside triphosphates were removed from the PCR products with the QIA Quick PCR Purification Kit (Qiagen). Amplified product was detected with an oligonucleotide probe specific for each virus labeled with horseradish peroxidase, using avidin-coated plates and hybridization buffer included in the commercially available Hexaplex assay (Prodesse). After hybridization at 50°C for 1 hour, the plates were washed, substrate added, the reaction stopped after 10 minutes, and optical density of each well measured at 450 nm on a spectrophotometer. The positive cutoff value was 3 SDs above the mean of water controls.

**Picornavirus**

The primers, composition of PCR master mix, amplification parameters, and probe for hybridization were detailed (adapted protocol from25). Reverse transcriptase-PCR with this method on log dilutions of a titered pool of type 39 rhinovirus demonstrated that ≥1 tissue culture infectious dose (TCID)_{50}/0.1 mL of sample was detected.

**Coronavirus**

The same PCR master mix25 was used for coronaviruses OC43 and 229E, with primers, probe, and amplification parameters for a sequence of the M gene of the 2 viruses as previously detailed. The sensitivity of the assay for 229E was 0.01 TCID_{50}/0.1 mL on the basis of titration in MRC-5 cells (human fetal lung fibroblast cell line).

**Parainfluenza 1–3, Influenza A and B**

The assay for these viruses was conducted by using Hexaplex reagents obtained from the manufacturer (Prodesse). Amplification of cDNA with a proprietary PCR mix containing primers for the viruses (Supermix, Prodesse) was performed by using cycling parameters detailed by Prodesse. Optical density at 450 nm of ≥0.4 was the positive cutoff, in accord with the manufacturer’s directions.
RSV

Primers, probe, and PCR cycling parameters were based on those described previously. The sensitivity of the assay determined by using a pool of RSV titrated in MRC-5 cells (kindly provided by Dr Ron Turner) was 0.28 TCID50/280 µL.

Adenovirus

PCR for adenovirus is based on the generic PCR developed by Echavarria and colleagues for detection of all clinically relevant adenovirus strains. The master mix for PCR contained 0.8 µM of each primer (biotinylated Hex 3 and Hex 4 primers) in 1× AmpliTaq Gold buffer with 2.7 mM MgCl2 and AmpliTaq Gold polymerase (Applied Biosystems, Foster City, CA). The amplification parameters were as detailed in a previous publication. The oligonucleotide probe Hex 30 labeled with horseradish peroxidase was used in the hybridization step. The sensitivity of PCR for adenovirus was <1 TCID50/0.1 µL on the basis of titration of an unnumbered adenovirus clinical isolate in A549 cells (kindly provided by Mike Ison, MD, MS).

Results for the 513 assays were cast as strings, and the number of reported assays was reduced to 425 independent assays by linking identical virus detections within a 20-day period but without an intermittent observation of “no detectable virus” as a single virus detection (assay), and by assigning multiple assays for the same CLI and/or OM string to the virus detected during that string or to no virus detected as applicable.

For each child, the CLI strings, virus strings, and the left and right OM strings were aligned on the time axis. First, CLI strings were examined for relationships to viruses, classified into 3 categories: CLIs with associated viruses (viruses immediately preceding or embedded within the string), CLIs with no viruses (containing embedded virus negative observations), and CLIs of unknown etiology (strings for which no virus samples were temporally associated), and counted. CLIs with associated viruses were labeled by the viral species, and CLIs that included both positive and negative virus recoveries were labeled by the recovered virus.

OM episodes (either unilateral or bilateral) were classified into 4 groups: OM as a CLI complication (episodes embedded within or immediately after the CLI, with primary etiology assigned to the associated virus, if any), OM as a complication of a viral infection but without evidence of a coexisting CLI (episode with a positive virus sample during the extant OM period), OM independent of a CLI/virus etiology (OM not associated with a CLI and for which embedded virus assays were negative), and OM of unknown etiology (episode without a coexisting CLI and no virus sampling during the episode duration).

For consistency, the summary format for data presentation was average ± SD, which is used throughout.

RESULTS

A total of 31 families were enrolled into the study in Year 1, and of these, 7 withdrew at some time before completing the study period. The reasons for withdrawal were that the study demands were overburdening to the family (n = 5) or the family relocated (n = 2). Consequently, the presented results focus on the data available for 60 children (3 black, 57 white) aged 4.4 ± 2.0 years (1.5–9.3 years) in the 24 families who completed the study and distributed as 2 children for 13 families, 3 children for 10 families, and 4 children in 1 family. Throughout the study period lasting from October 1 to April 30, a total of 513 nasal samples were collected and analyzed by PCR, yielding an average of 9.6 ± 2.5 (range: 2–16) samples per child (425 independent assays, see “Methods”); a total of 1513 bilateral pneumatic otoscopic examinations were done for an average of 25.2 ± 4.0 (range: 9–32) bilateral examinations per child, and the daily diaries were complete for 12 296 child days, with an average of 205 ± 6 (range: 197–213) completed days per child. These results indicate excellent compliance with the planned otoscopic examinations and other data collections that is attributable primarily to the “in-home” visits made by study personnel.

Two-thousand seventy-seven of all child days (17%) were assigned by algorithm to an illness day, and the average cold burden when defined by the percentage of observation days assigned to an illness day was 17 ± 15% (range: 0%–68%) per child. These days were linked as 199 CLIs with an average rate of 3.3 ± 2.0 (range: 0–7) CLIs per child. Nasal samples were not collected during the CLI in 71 cases (36%) because of the lack of free secretions in children who refused to have nasal washing done, unavailability of the family during the CLI period (vacations, holidays), or the failure of the mother to assign and/or report a cold to study personnel during times when a CLI was identified retrospectively by the algorithm. Of the 128 CLIs with secretion sampling (64%), a virus was identified in 94 (73%). For individual children, virus was detected during an average of 2.1 ± 1.3 (range: 0–5) of the sampled CLIs for an average detection rate of 77 ± 27%. The distribution of viruses and virus combinations temporally associated with a CLI is reported in Table 1. As expected, for those CLIs with identified viral etiologies, rhinovirus was associated with >50% of the episodes.

Of the 297 independent assays for virus performed during non-CLI periods, 54 (18.2%) were PCR positive for a virus or virus combination. Species distributions for these viruses are reported in Table 1. For all subjects, the average percentage rate of positive virus recovery given a non-CLI period was 16 ± 17% (range: 0%–67%). The relative frequency distribution of the viruses and virus combinations is similar to that for the CLIs, with rhinovirus most frequently identified, followed by RSV and influenza, and then adenovirus, parainfluenza, and...
coronavirus. Two-hundred forty-three specimens recovered during the non-CLI periods yielded no virus (81.8%) and, for individual subjects, the average frequency of negative virus recovery in the absence of a CLI was 84 ± 17% (range: 33%–100%).

Overall, there were 93 diagnosed OM episodes (either unilateral or bilateral). The average number of episodes per subject was 1.6 ± 1.3 (range: 0–4). Table 2 shows the distribution of viruses associated with these episodes when a CLI was diagnosed and when virus was identified in the absence of a CLI. Sixty-five of all OM diagnoses (70%) were related to a concurrent CLI (irrespective of concurrent virus recovery). Of these, secretion samples were not obtained in 10 cases (15%) and, in an additional 12 cases (18%), a virus was not identified despite inclusive sampling. The relative ordering of viruses by frequency of recovery for OM was not different from that presented for CLIs (see Table 1).

More interesting is the relative frequency of OM during non-CLI periods but when viruses were identified in the nasal sample. There, an additional 18 OM episodes (19%) were concurrent with a virus isolated from the nose and nasopharynx, 6 OM episodes (6%) were not associated with viruses despite sampling during the OM episode, and 4 OM episodes (4%) could not be assigned because secretion samples were not collected during the episode. As with CLIs, the majority of these OM episodes were associated with rhinoviruses and a minority with adenovirus and coronavirus, but influenza and RSV were conspicuously absent.

For this population, the OM complication rate given a CLI was estimated at 65 OM episodes per 199 CLIs or

| TABLE 1 | Distribution of Recovered Virus for CLI and Non-CLI Periods |
|-----------------|-----------------|-----------------|
| Virus 1 | Virus 1+ | CLI | No CLI |
| | | | |
| n | Frequency, % | n | Frequency, % |
| Rhinovirus | — | 52 | 55.3 | 36 | 66.7 |
| Rhinovirus | Coronavirus | 3 | 3.2 | 3 | 5.6 |
| Rhinovirus | RSV | 3 | 3.2 | 2 | 3.7 |
| Rhinovirus | Influenza | 2 | 2.1 | 1 | 1.9 |
| Rhinovirus | Adenovirus | 1 | 1.1 | 3 | 5.6 |
| Influenza | — | 8 | 8.5 | 1 | 1.9 |
| Influenza | RSV | 2 | 2.1 | 0 | 0.0 |
| RSV | — | 6 | 6.4 | 2 | 3.7 |
| Adenovirus | — | 5 | 5.3 | 2 | 3.7 |
| Adenovirus | Coronavirus | 1 | 1.1 | 0 | 0.0 |
| Parainfluenza | — | 6 | 6.4 | 1 | 1.9 |
| Coronavirus | — | 5 | 5.3 | 3 | 5.6 |
| Total with virus | — | 94 | 73.4b | 54 | 18.2 |
| No virus recovery | — | 34 | 26.6b | 243 | 81.8 |
| No secretion sample | — | 71 | — | NA | — |
| Total | — | 199 | — | 297 | — |

NA indicates not applicable.

| TABLE 2 | Distribution of Recovered Virus for OM Episodes According to CLI Presence/Absence |
|-----------------|-----------------|-----------------|
| Virus 1 | Virus 2 | OM With CLI | OM Without CLI |
| | | Count | % | Count | % |
| Rhinovirus | — | 20 | 30.8 | 14 | 50.0 |
| Rhinovirus | RSV | 0 | 0.0 | 1 | 3.6 |
| Rhinovirus | Adenovirus | 0 | 0.0 | 1 | 3.6 |
| RSV | — | 6 | 9.2 | 0 | 0.0 |
| RSV | Influenza | 2 | 3.1 | 0 | 0.0 |
| Influenza | — | 5 | 7.7 | 0 | 0.0 |
| Parainfluenza | — | 3 | 4.6 | 0 | 0.0 |
| Adenovirus | — | 3 | 4.6 | 1 | 3.6 |
| Adenovirus | Coronavirus | 1 | 1.5 | 0 | 0.0 |
| Coronavirus | — | 3 | 4.6 | 1 | 3.6 |
| None identifieda | — | 12 | 18.5 | 6 | 21.4 |
| Unknownb | — | 10 | 15.4 | 4 | 14.3 |
| Total | — | 65 | 100.0 | 28 | 100.0 |

a Virus-negative samples obtained during OM periods.

b OM episodes with no inclusive secretion samples.
33%. Of the 79 OM episodes with concurrent nasal samples, ≥1 virus was isolated in 61 (77.2%). Of the 94 CLIs with virus recovery, 43 (45.7%) were complicated by a new OM episode. Finally, a new OM episode was associated with 18 (33.3%) of the 54 times virus was recovered from a nasal sample in the absence of a CLI.

DISCUSSION
The results present a set of temporal coincidences among virus isolations from the upper respiratory tract, CLI expression, and OM as an otologic complication that suggest a causal relationship. Past studies in human volunteers support this interpretation by showing that experimental infection of the nose and nasopharynx with rhinovirus, influenza A, or RSV provokes a variable expression of signs/symptoms of illness that qualifies as a CLI episode in approximately two thirds of those exposed and otologic complications including OM in some but not all subjects.1,3,8 Other analyses of those data showed that illness and otologic expressions are decoupled; for example, a CLI is not a prerequisite to otologic complications, and the 2 show differential responses to prechallenge antibody titer and antiviral therapy.1,18,20,21 Because upper respiratory viruses are recovered from the nose and nasopharynx of infants, children and adults in the absence of overt illness,12–25 these results suggested to us that the OM complication rate of a vURI when based on a concurrent CLI underestimates the true rate, and similarly, that the percentage of total OM episodes explained by a vURI is much higher than previously reported.

The species distribution of viruses during CLIs in this study is similar to that reported previously, with rhinoviruses accounting for most (~60%) illnesses, followed by influenza A and RSV at 10% and then adenovirus, coronavirus, and parainfluenza at ~5% each.1,3,8 It is interesting that a similar relative frequency distribution for these viruses was documented for non-CLI periods with positive virus recovery. It is expected that this distribution reflects the proportional sampling over the entire “cold season” and that the frequency of virus recovery would be biased to a particular species by concentrated sampling during an endemic period.

Approximately 70% of OM episodes were temporally associated with a CLI, and where ascertainment was possible, there was no evidence of skewing in favor of specific viruses. Rather, the CLIs complicated by OM showed the same relative distribution of viruses as the source population from which they were drawn. This observation stands in contrast to some previous reports that specified a higher than anticipated frequency of OM episodes for a particular virus species.1,2,4,6 Because our result agrees well with more comprehensive, cross-seasonal surveys,1,3,8 we hypothesize that those reports were biased by factors that included dependence on signs/symptoms of illness for OM identification, the source population (eg, hospitalized children), and study season. Finally, the study data reject our null hypotheses. Both hypotheses require that viral recovery during non-CLI periods not be associated with OM, a formulation that is inconsistent with our results. Rather, OM occurs as a complication of the presence of virus in the nose and/or nasopharynx irrespective of CLI presentation. Moreover, although the sample set is small, there does not seem to be a clear preferential ordering as to which viruses are associated with OM in the absence of illness. In that regard, our assessment of the presence or absence of a CLI was made by using an algorithm, and it is possible that the code set was insensitive to sign patterns specific to a given virus or to alternative sign presentations. To test this possibility, a 20-day period centered on each observation of a non-CLI-associated positive viral recovery was examined for concurrent signs. The results were consistent with a typically null set or a sporadic expression of signs that could not be clearly related to the virus (data not shown).

CONCLUSIONS
Seventy percent of all OM episodes were attributable to a concurrent CLI episode, and, in this population of older children, OM was a complication in 33% of all CLI episodes. The primary etiologic agent associated with CLI and OM episodes was rhinovirus. Respiratory viruses were recovered during a large number of non-CLI periods, and these too were associated with the development of OM. More recently, other respiratory viruses such as metapneumovirus and bocavirus were identified as causing CLIs and precipitating OM, but their presence was not assayed in our study because the requisite techniques are not yet available in our laboratories.33,34 It is possible that these and other viruses were causative in those CLIs and OM episodes where our assay panel failed to identify a viral etiology. We conclude that most OM episodes are a complication of a vURI and that past age-adjusted estimates of OM as a vURI complication are biased downwardly by their dependence on CLI presentation.

ACKNOWLEDGMENTS
This study was supported, in part, by National Institutes of Health grant DC005832.

We thank Kathleen Ashe for assisting with the virology assays; Margaretha L. Casselbrant, Harriette Wheatley, Kathy Tekely, and Ellen Reynolds for assisting with otoscopic examinations and sample procurement; and Julianne Banks and James T. Seroky for assisting with data entry.

REFERENCES


Immunogenicity and Safety of Intradermal Influenza Immunization at a Reduced Dose in Healthy Children

Susan S. Chiu, MD, FAAP, J. S. Malik Peiris, DPhil, FRCPath, FRSA, Kwok H. Chan, PhD, Wilfred Hing Sang Wong, MMedSc, Yu Lung Lau, MD

Departments of Pediatrics and Adolescent Medicine and Microbiology, University of Hong Kong, Hong Kong, China; Pasteur Research Centre, Hong Kong University, Hong Kong, China

Financial Disclosure: Dr. Lau has conducted clinical trials for Wyeth, GlaxoSmithKline, Medimmune, and Merck Sharp and Dohme (Asia) Limited and is a member of the Steering Committee for Prevention and Control of Infectious Diseases in Asia for GlaxoSmithKline. The other authors have indicated they have no financial relationships relevant to this article to disclose.

ABSTRACT

OBJECTIVES. We conducted this study to test the hypothesis that intradermal influenza vaccination at one fifth of a standard dose elicits comparable immunogenicity to full-dose intramuscular vaccination in children.

PATIENTS AND METHODS. We conducted a randomized, open-label study in 112 healthy children aged 3 to <18 years to compare the immunogenicity and safety of intradermal vaccination at one fifth of a dose with standard intramuscular vaccination. Analyses of hemagglutination inhibition antibody titers to each antigen in each group included geometric mean titers before and 21 days after vaccination, fold increase in geometric mean titers after vaccination, seroprotection rate, and seroconversion rate.

RESULTS. The mean age of the subjects was 10.11 ± 4.04 years in the intradermal vaccination group and 10.57 ± 3.91 years in the intramuscular group. Intradermal vaccination was safe. Induration and mild erythema at the injection site were reported at 25% and 57%, respectively, in the intradermal group. Fold increase of geometric mean titers against influenza A/Caledonia was robust in both groups (11.1-fold and 12.9-fold increase in the intramuscular and intradermal groups, respectively), whereas that for B/Shandong was more modest (4.3–4.4). Both approaches elicited very high geometric mean titers against influenza A/Panama: 1360.5 and 893.9 for the intramuscular and intradermal groups, respectively, but because the prevaccination antibody titers were high, the fold increase of geometric mean titers was only 4.5 and 2.6, respectively.

CONCLUSION. The immunogenicity of one fifth of a dose of influenza vaccine delivered by the intradermal route is comparable to the standard-dose intramuscular vaccination in children as young as 3 years of age.
Seasonal interpandemic influenza infection has serious health impacts worldwide, and vaccination is an important public health intervention.\(^1\) Vaccination is also a key intervention in containing an influenza pandemic.\(^6\) Even with well-established vaccine production for interpandemic influenza, shortage of supply sometimes occurs. Therefore, there is justified concern that when large populations need to be vaccinated within a short period of time, demand may exceed production capacity.

Intradermal vaccination exploits the abundance of antigen-presenting cells (macrophages and dendritic cells) that allow a robust immune response to be elicited with a small dose of antigen delivered directly to the skin, which may be a solution to vaccine shortage. Intradermal vaccination is not new. Bacille Calmette-Guerin vaccination, which is widely practiced in many parts of the world, is routinely performed by the intradermal route. Studies exploring intradermal vaccination were conducted with hepatitis B, rabies, and in influenza vaccines as early as the 1940s.\(^10\)-\(^16\) However, these early studies either had no control group or had control groups using the subcutaneous method of administration, were conducted with monovalent or bivalent vaccines with antigen dosages not directly comparable to those available now, lack a clear definition of response in at least 1 study, or used assays that were crude when compared with what is available today. Two recent articles reported on the effectiveness of intradermal influenza vaccination at a reduced dose in adults.\(^17\)-\(^18\) It is important to investigate the possibility of influenza vaccination using the intradermal approach in children.

METHODS
Study Design
This was an open-label, randomized study to compare the safety and immune responses of intramuscular administration of an inactivated influenza vaccine with intradermal administration of the same vaccine at one fifth of a dose in healthy children 3 to \(<18\) years of age. The hypothesis was that intradermal vaccination with one fifth of a standard dose of influenza vaccine elicited comparable immunogenicity as full dose of intramuscular vaccination of the same vaccine. The study protocol was approved by the joint institutional review board of the University of Hong Kong and Queen Mary Hospital (Hong Kong) and was in accordance with the Helsinki Declaration of 1975, as revised in 1983. Subjects were patients who were admitted to the hospital for an acute illness or followed in our outpatient clinic. Written informed consent was obtained from parents and the older children before enrollment. Subjects were considered ineligible if they had an underlying condition that rendered them at risk for influenza complications or a need for regular medication. For children between 3 and \(<9\) years of age, only those who had previous influenza vaccination were recruited to avoid the need for a second dose of immunization. Subjects were matched by age, and a computer-generated randomization list with a block size of 4 was used to assign study subjects to receive an intramuscular dose of 0.5 mL of inactivated trivalent influenza vaccine, or an intradermal dose of 0.1 mL of the same vaccine. Randomization was performed by study personnel immediately before vaccination. The randomization assignment was blinded to the laboratory investigators. All subjects had height and weight measured and the history of influenza vaccination in the previous year elicited. The children were vaccinated in October and November 2005, before the influenza season in Hong Kong that usually peaked in January or February.

Vaccination
The influenza vaccine (Fluarix; GlaxoSmithKline Biologicals, Rixensart, Belgium) we used contained purified surface antigen equivalent to the influenza types and subtypes recommended by the World Health Organization for the 2005–2006 season: A/New Caledonia/20/99 (H1N1)-like strain; A/California/7/2004 (H3N2)-like strain, and B/Shanghai/361/2002-like strain. Each 0.5-mL dose contains at least 15 \(\mu\)g of hemagglutinin antigen per recommended strain.

Intramuscular injection was performed in the deltoid area according to standard procedure. Intradermal injection was performed using a 1-mL syringe calibrated in hundredths (Becton Dickinson, Franklin Lakes, NJ). One research nurse performed all the injections at each subject’s deltoid area (Fig 1). The needle was inserted at a 15° angle to the skin according to standard intradermal injection techniques.\(^18\) The resulting wheal was graded by the research nurse as 0 (no wheal), 1 (wheal size of 1–2 mm in diameter), 2 (3–5 mm in diameter), or 3 (>5 mm in diameter).

FIGURE 1
A, Intradermal vaccination using the 1-hand technique with the help of a caretaker in a 3-year-old boy during the study. B, Intradermal vaccination using the 1-hand technique.
Subjects were observed for 15 to 30 minutes after vaccination for acute reactions. Caretakers were given a diary card for recording of reactions for 3 days. Solicited reactions were fever, malaise, shivering, erythema, induration, and bruising of the injection site of reactions were fever, malaise, shivering, erythema, induration, and bruising of the injection site of. Adverse effects were graded according to severity. A mild adverse event was one in which the symptoms were easily tolerated, a moderate reaction caused interference with usual activities, and a serious one resulted in inability to perform usual activities. Research personnel retrieved the adverse events information on day 4 by telephone. Parents were asked to bring back the diary card for cross-checking when the subjects returned for the postvaccination blood draw.

Serologic Studies and Outcome Measures

Serum samples were drawn on day 0 (before vaccination) and day 21 from each subject. The paired samples were tested together by hemagglutination inhibition (HAI) using reference antigens provided as part of the World Health Organization kit provided by the World Health Organization Influenza Collaborating Centre, Centers for Disease Control and Prevention (Atlanta, GA). The sera were treated with receptor-destroying enzyme (RDE) (1:3) at 37°C overnight to remove non-specific inhibitors, and residual RDE was destroyed by heat inactivation at 56°C for 30 minutes. Serial twofold dilutions of RDE-treated serum (1 in 10) were titrated in a 96-well microtiter plate against 4 hemagglutinin units of reference antigens (H1N1, H3N1, and B/Shanghai/361/2002-like) using 0.25% turkey erythrocytes.

The following analyses were conducted on the antibody titers to each antigen obtained in the 2 groups of patients: geometric mean titers (GMTs) before and at 21 days after vaccination; fold increase in GMT (calculated as the mean of the ratio of titer before and 21 days after vaccination); the seroprotection rate (defined as the percentage of patients with antibody titers ≥40); and the seroconversion rate (calculated as the percentage of subjects with a prevaccination antibody titer of <40 who developed a fourfold rise or a titer of ≥40 after vaccination). Those with a fourfold increase in antibody titer were determined for the intramuscular and intradermal vaccine groups overall, with subset analyses for those whose prevaccine antibody titer was <40 and for those ≥40.

Statistical Analysis

The 2 methods of vaccine administration are considered equivalent if the 95% confidence interval (CI) for the ratio of geometric mean titers (as a percentage) is sufficiently narrow and lies within a range of 80% to 120%. A sample size of 50 subjects per group would have 80% statistical power to identify a difference between the 2 methods using a 2-sided test and a type 1 error rate of 5%, to detect a significant difference of a minimum difference of a ratio of 2.05 in the geometric mean titer and a ratio of 2.33 in the fold increase between the 2 groups. Fifty-six subjects were recruited for each group to take into account an attrition rate of 12%. For the intention-to-treat analysis, missing values of convalescent serology would be assigned the mean value of the group. Pearson correlation was used to detect correlation between individual titers in the intradermal group with age, induration size and BMI. Adverse events for the 2 groups were compared by means of the χ² test.

RESULTS

Fifty-six subjects each were assigned randomly to receive either intramuscular or intradermal influenza vaccination. Two children in the intramuscular group failed to return for convalescent blood taking. The demographics of the subjects are shown in Table 1. They were comparable in age, male to female ratio, BMI, and history of influenza vaccine in the previous year. When using the 1-hand injection technique, the intradermal vaccination did not pose significant problems, even in young children. Median score of the intradermal lesions was 2 (3- to 5-mm wheal), with a range of 1 to 3.

Reactogenicity

Complete reactogenicity data were available from all subjects. Both methods of influenza vaccination were well tolerated (Table 2). Not surprisingly, induration was reported more in the intradermal group, but disappeared by the next day in the majority. Mild erythema was reported in 57.1% of the intradermal group, significantly more frequent than in the intramuscular group (3.6%). There was no other significant difference of either local or systemic adverse effects in both groups.

Immunogenicity

Of the 112 subjects recruited and vaccinated, 110 returned for convalescent blood draw. Two subjects in the intramuscular group failed to return for convalescent blood draw. The demographics of children randomly assigned to receive either 0.1 mL of Influenza Vaccine by Intramuscular Injection or 0.5 mL of Influenza Vaccine by Intradermal Injection are shown in Table 1. They were comparable in age, male to female ratio, BMI, and history of influenza vaccine in the previous year. When using the 1-hand injection technique, the intradermal vaccination did not pose significant problems, even in young children. Median score of the intradermal lesions was 2 (3- to 5-mm wheal), with a range of 1 to 3.

TABLE 1 Demographics of Children Randomly Assigned to Receive Either 0.1 mL of Influenza Vaccine by Intramuscular Injection or 0.5 mL of Influenza Vaccine by Intradermal Injection

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Intradermal Group (n = 56)</th>
<th>Intramuscular Group (n = 56)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age, y</td>
<td>10.11 ± 4.04</td>
<td>10.57 ± 3.91</td>
<td>.546</td>
</tr>
<tr>
<td>Age range, y</td>
<td>3.2–17.4</td>
<td>3.1–17.2</td>
<td>1</td>
</tr>
<tr>
<td>Male/female ratio</td>
<td>30.26</td>
<td>30.26</td>
<td>.442</td>
</tr>
<tr>
<td>Mean BMI</td>
<td>18.3 ± 3.3</td>
<td>17.8 ± 4</td>
<td>.472</td>
</tr>
<tr>
<td>Range of BMI</td>
<td>13.7–30.2</td>
<td>12.8–30.9</td>
<td>.393</td>
</tr>
<tr>
<td>Influenza vaccine in past year, yes, %</td>
<td>17 (30.4)</td>
<td>13 (23.2)</td>
<td>.493</td>
</tr>
<tr>
<td>Age, n/N (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;9 y</td>
<td>12/16 (75)</td>
<td>6/15 (40)</td>
<td>.467</td>
</tr>
<tr>
<td>≥9 y</td>
<td>5/40 (12.5)</td>
<td>7/41 (17)</td>
<td>.849</td>
</tr>
</tbody>
</table>
return for convalescent serology and were assigned the mean serology titers of the group. Convalescent blood was drawn at a mean of 23.7 ± 3.1 day and 24.4 ± 5.9 days after vaccination in the intradermal and intramuscular group, respectively (P = not significant). All blood was drawn before the onset of influenza season in January 2006.

There was no statistically significant difference in seroconversion rate and seroprotection rate after vaccination between subjects in the 2 groups against any of the 3 antigens (Table 3). High postvaccination GMT were elicited in both groups against all 3 antigens, especially for the 2 influenza A antigens. There was no correlation of all individual antibody titers with age, BMI, or history of influenza vaccine in the previous year. There was also no correlation between antibody titers and induration size in the intradermal group (data not shown). There was no significant difference between the intramuscular and intradermal groups in the fold increase in GMT antibody titer for H1N1, at 11.1- and 12.9-fold increase respective, or B/Shandong, at 4.4- and 4.3-fold increase, respectively. However, there was a significant difference in fold increase of GMT against H3N2, with a fold rise of 4.5 in the intramuscular group compared with that of 2.6 in the intradermal group (P = .005). Despite this difference, both approaches elicited very high GMT: 1360.5 and 893.9 for the intramuscular and intradermal groups, respectively. For all those with prevaccination antibody titer of <40 to any of the 3 antigens (n = 68), there was no difference in fold-rise in GMT for all antigens combined: 12.5-fold in the intramuscular group and 17.9-fold in the intradermal group, suggesting that the response in seronegative subjects using either method of vaccine administration was equally robust.

A significant proportion of subjects in both groups had at least a fourfold increase in titers against the 3 antigens. There was, again, a significant difference in the proportion of children with at least a fourfold rise in GMT against H3N2 between the intramuscular group (66%) and the intradermal group (43%). Those with a lower prevaccination titer were more likely to respond with at least a fourfold rise of titer. A higher percentage of subjects in either group with prevaccination titers of <40 against any of the 3 antigens had at least fourfold increase in GMT when compared with those who had prevaccination titers of ≥40 (Table 3). More subjects with high prevaccination titers had a modest or no increase in titer after vaccination in the intradermal group. Twenty-seven subjects (55%) in the intradermal group with prevaccination titers against H3N2 ≥1: 160 had less than a fourfold rise in titer, as compared with 16 such subjects (33.3%) in the intramuscular group (P = .04).

To validate the power calculation of our sample size, we calculated, with the fold increase defined relative to day 0, the standard deviations of GMT and the mean log10 ratio of GMT to be 0.66 and 0.764, respectively, with little variation according to strain. The sample size of 56 in each group, therefore, had 80% statistical power to identify a difference between the 2 methods using a 2-sided test and a type 1 error rate of 5%, for a ratio of 2.07 in GMT and a ratio of 2.52 in the fold increase.

**DISCUSSION**

Intradermal administration of influenza vaccine in children was safe and immunogenic. Local reaction was more common in subjects who received the intradermal injection. This finding is similar to that reported in a study of intradermal rabies vaccine in children.15 We did not find any difference in systemic reactions between the 2 groups.

Very high HAI GMTs against all 3 antigens of the influenza vaccine were elicited by the intradermal route by using one fifth of the standard dose. These titers were much higher than those reported in similar adult studies.17,18 Without direct comparison, it is not known whether intradermal influenza vaccination in children is in fact more immunogenic than in adults. However, available data show that young skin is superior to old skin in resting Langerhans cell numbers and migration response after intradermal injection.21 Because H3N2 has been in circulation worldwide since 1968 and H1N1 since 1977, and young children with previous influenza vaccination were recruited by design, it is not surprising that the majority of children with a mean age of 10 years in this study already had HAI titers of ≥1:40 against both antigens. The protective and high prevaccination GMT supports the notion that children in Hong Kong are heavily influenza experienced. It was suggested that in a partially seropositive population like that in the current study, fold increase in titers tends to underestimate vaccine immunogenicity.22 This, rather than the route of administration, may partly explain the statistical signif-

<table>
<thead>
<tr>
<th>Symptoms</th>
<th>Intramuscular Group (n = 56), %</th>
<th>Intradermal Group (n = 56), %</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Induration (mild)</td>
<td>5.4 25</td>
<td>5.4 25</td>
<td>.007</td>
</tr>
<tr>
<td>Erythema (mild)</td>
<td>3.6 57.1</td>
<td>4.5 25</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Ecchymosis &gt;0.5 cm</td>
<td>0 5.4</td>
<td>0 5.4</td>
<td>NS</td>
</tr>
<tr>
<td>Itchiness around injection site</td>
<td>0 5.4</td>
<td>0 5.4</td>
<td>NS</td>
</tr>
<tr>
<td>Pain around injection site</td>
<td>5.4 1.8</td>
<td>5.4 1.8</td>
<td>NS</td>
</tr>
<tr>
<td>Malaise (overall)</td>
<td>23.2 26.8</td>
<td>23.2 26.8</td>
<td>NS</td>
</tr>
<tr>
<td>Mild</td>
<td>16.1 17.9</td>
<td>16.1 17.9</td>
<td>NS</td>
</tr>
<tr>
<td>Moderate</td>
<td>1.8 7.1</td>
<td>1.8 7.1</td>
<td>NS</td>
</tr>
<tr>
<td>Severe</td>
<td>5.4 1.8</td>
<td>5.4 1.8</td>
<td>NS</td>
</tr>
<tr>
<td>Shivering/chills</td>
<td>5.4 3.6</td>
<td>5.4 3.6</td>
<td>NS</td>
</tr>
<tr>
<td>Fever &gt;38°C</td>
<td>7.1 5.4</td>
<td>7.1 5.4</td>
<td>NS</td>
</tr>
<tr>
<td>Headache</td>
<td>1.8 3.6</td>
<td>1.8 3.6</td>
<td>NS</td>
</tr>
<tr>
<td>Cough</td>
<td>3.6 3.6</td>
<td>3.6 3.6</td>
<td>NS</td>
</tr>
<tr>
<td>Hoarseness of voice</td>
<td>1.8 0</td>
<td>1.8 0</td>
<td>NS</td>
</tr>
<tr>
<td>Muscle pain</td>
<td>0 1.8</td>
<td>0 1.8</td>
<td>NS</td>
</tr>
</tbody>
</table>

NS indicates not significant.
icant difference in GMT fold increase against H3N2 observed. This hypothesis is supported by (1) the comparable fold rises against H1N1 and influenza B/Shandong, against which subjects from both groups had a lower prevaccination GMT and were less likely to have an antibody level \( < 1:40 \), (2) a higher percentage of at least fourfold rise in titer in both groups in those with seronegative prevaccination titers, and (3) a robust and equivalent fold increase for seronegative prevaccination titers against all 3 antigens in both groups. In rabies vaccination of children, there has been concern that antibodies elicited by intradermal vaccination were of lower titer, and may not be as durable.15 With very high postvaccination GMT, it is unlikely that the difference in postvaccination GMT against H3N2 between the 2 groups would be clinically significant. Moreover, with influenza vaccine being given annually, the long-term durability of antibodies is less relevant.

The skin has been the target of immunization for many years, and there is recent renewal of interest in exploiting this immune organ. Studies performed in animals and humans have had promising results.23–25 Epidermal vaccination that includes transcutaneous immunization or needle-free jet injectors targets the epidermis that is abundant in Langerhans cells, and epidermal and dermal vaccination that uses gene gun or electroporation technology targets both layers of the skin that contains Langerhans cells as well as dendritic cells.26 These methods of vaccination require special preparations and devices. Intradermal vaccination targets the dermis that is abundant in dendritic cells and can be performed by a simple needle and syringe using existing vaccine prepa-

### TABLE 3 Strain-Specific HAI Result for the 3 Antigenic Components of the Trivalent Influenza Vaccine

<table>
<thead>
<tr>
<th></th>
<th>Intramuscular Group</th>
<th>Intradermal Group</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geometric mean (95% CI)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1N1 Day 0</td>
<td>72.6 (50.6–103.9)</td>
<td>49.9 (34.7–71.9)</td>
<td>NS</td>
</tr>
<tr>
<td>Day 21</td>
<td>803.7 (560.8–1152.1)</td>
<td>647.9 (435.2–964.8)</td>
<td>NS</td>
</tr>
<tr>
<td>H2N2 Day 0</td>
<td>302.8 (219.6–417.5)</td>
<td>340.4 (249.9–463.8)</td>
<td>NS</td>
</tr>
<tr>
<td>Day 21</td>
<td>1360.5 (1041.5–1777.3)</td>
<td>893.9 (685.9–1165.1)</td>
<td></td>
</tr>
<tr>
<td>B/Shanghai Day 0</td>
<td>90.9 (65.1–126.8)</td>
<td>88.3 (62.9–123.9)</td>
<td>NS</td>
</tr>
<tr>
<td>Day 21</td>
<td>400 (272.4–587.5)</td>
<td>385.3 (255.6–580.8)</td>
<td>NS</td>
</tr>
<tr>
<td>Seroconversion rate, ( n/N ) (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1N1 Day 0</td>
<td>14/15 (93)</td>
<td>21/22 (95)</td>
<td>NS</td>
</tr>
<tr>
<td>Day 21</td>
<td>15/15 (100)</td>
<td>22/22 (100)</td>
<td>NS</td>
</tr>
<tr>
<td>H3N2 Day 0</td>
<td>9/13 (69)</td>
<td>9/11 (82)</td>
<td>NS</td>
</tr>
<tr>
<td>Day 21</td>
<td>4.2/6 (75)</td>
<td>42/56 (75)</td>
<td>NS</td>
</tr>
<tr>
<td>B/Shanghai Day 0</td>
<td>37/56 (66%)</td>
<td>24/56 (43%)</td>
<td>0.05</td>
</tr>
<tr>
<td>Day 21</td>
<td>19/56 (34)</td>
<td>29/56 (52)</td>
<td>NS</td>
</tr>
<tr>
<td>Fold increase in geometric mean titer (95% CI)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1N1 Day 0</td>
<td>11.1 (7.2–17.1)</td>
<td>12.9 (7.9–21.4)</td>
<td>NS</td>
</tr>
<tr>
<td>Day 21</td>
<td>4.3 (3.4–5.9)</td>
<td>2.6 (2.1–3.4)</td>
<td>0.05</td>
</tr>
<tr>
<td>H3N2 Day 0</td>
<td>4.4 (3.3–5.8)</td>
<td>4.3 (3.2–5.8)</td>
<td>NS</td>
</tr>
<tr>
<td>Day 21</td>
<td>4.0/6 (75)</td>
<td>42/56 (75)</td>
<td>NS</td>
</tr>
<tr>
<td>B/Shanghai Day 0</td>
<td>37/56 (66%)</td>
<td>24/56 (43%)</td>
<td>0.03</td>
</tr>
<tr>
<td>Day 21</td>
<td>19/56 (34)</td>
<td>29/56 (52)</td>
<td>NS</td>
</tr>
<tr>
<td>Fold increase in subjects with prevaccination ( &lt; 40 ), ( n/N ) (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1N1 Day 0</td>
<td>14/15 (93)</td>
<td>21/22 (95)</td>
<td>NS</td>
</tr>
<tr>
<td>Day 21</td>
<td>3/3 (100)</td>
<td>3/3 (100)</td>
<td>NS</td>
</tr>
<tr>
<td>H3N2 Day 0</td>
<td>10/20 (50)</td>
<td>20/27 (74)</td>
<td>NS</td>
</tr>
<tr>
<td>Day 21</td>
<td>28/41 (68.3)</td>
<td>21/34 (61.8)</td>
<td>NS</td>
</tr>
<tr>
<td>B/Shanghai Day 0</td>
<td>14/15 (93)</td>
<td>21/22 (95)</td>
<td>NS</td>
</tr>
<tr>
<td>Day 21</td>
<td>3/3 (100)</td>
<td>3/3 (100)</td>
<td>NS</td>
</tr>
<tr>
<td>Fold increase in subjects with prevaccination ( \geq 40 ), ( n/N ) (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1N1 Day 0</td>
<td>34/53 (64.2%)</td>
<td>21/53 (39.6%)</td>
<td>0.02</td>
</tr>
<tr>
<td>Day 21</td>
<td>9/36 (25)</td>
<td>9/29 (31)</td>
<td>NS</td>
</tr>
</tbody>
</table>

CI indicates confidence interval; NS, not significant.

\* Using unpaired \( t \) test.

\( ^{a} \) Using \( \chi^2 \) test.
ration. However, in general, intradermal injection requires more skill than intramuscular injection. Performing intradermal vaccination in struggling young children is more challenging than in obliging small infants. The current study has demonstrated that intradermal injection can be performed safely and effectively even in children as young as 3 years of age using the 1-hand injection technique, which can be easily mastered using standard tuberculin syringes and can also be performed effectively in overweight children with more subcutaneous fat. This has important implication for use in resource poor countries, because sophisticated devices are likely to mean increased costs.

There are limitations in our study. Our subjects were heavily influenza experienced. Although this situation reflects reality, where a proportion of the population is likely to be seropositive to the strains in the vaccine, this will not hold true regarding the pandemic strain. All the intradermal injections were administered by 1 experienced person. A larger variation of results may be expected when many persons give the vaccine. We designed the experiment to obtain the convalescent blood samples 21 days after vaccination. Ideally, another sampling on day 42 postvaccination would have provided additional information on antibody kinetics. However, repeated blood sampling is generally not well received by parents of healthy young children. Future studies should involve children younger than 3 years of age, children seronegative to the vaccine strains, children with various immunocompromised states, comparing 1 intradermal dose with the intramuscular 2-dose regimen in young children, multiple vaccine administrators, longer follow-up, and studies to include efficacy to ascertain that antibodies elicited by different routes of administration are comparable functionally.

This study has provided encouraging data that intradermal injection can elicit comparable immune response against influenza at a much reduced dose in healthy children as young as 3 years of age and is a useful strategy against vaccine shortage. Pandemic vaccines with avian hemagglutinins, especially the nonadjuvanted H5N1 vaccine, are known to be poor immunogens. It is not known whether the intradermal route of vaccination will offer any advantages. A recent report of intradermal administration of a subvirion H5N1 vaccine was not immunogenic up to 3 doses in healthy adults. We are currently investigating the cellular response of our subjects to better define the role of intradermal influenza vaccination.

**ACKNOWLEDGMENTS**

This work was supported by Research Grants Council of Hong Kong grant HKU 7396/03M, Ellison Medical Foundation grant 1D-1A-0036-02, and the Vice Chancellors Development Fund (University of Hong Kong).

**REFERENCES**


We thank the children and parents who participated in the study, our dedicated team of research staff, and Winnie Lau, who single-handedly performed all the intradermal injections.


28. Patel SM, Atmar RL, Sahly H, Cate TR, Keitel WA. Randomized, open-label, phase 1 clinical trial comparing the safety, reactogenicity, and immunogenicity of booster immunization with inactivated influenza A/H5N1 vaccine administered by the intradermal (ID) or the intramuscular (IIM) route among healthy adults [abstract]. Presented at: IDSA annual meeting; October 12–15, 2006; Toronto, Ontario, Canada. Abstract LB-5, 64

---

**YALE ON $0 A DAY**

“Getting into college may be tougher than it used to be. But top schools are offering a growing number of courses free online. Following the lead of the Massachusetts Institute of Technology and other highly competitive schools, more institutions are posting online everything from lecture notes to sample tests, and even making audio and video files of actual lectures publicly available. The sites attract anywhere from thousands to more than one million unique visitors each month. . . . MIT’s pioneering ‘OpenCourseWare’ program, which was launched in 2003, posts the syllabus and class notes for more than 1500 courses online for anyone who wants them. By this November, it aims to publish materials from virtually all 1800 of its courses across all its schools. . . . Starting last fall, the University of Notre Dame in South Bend, Ind., began offering eight courses. . . . Yale University, meanwhile, has announced it will produce digital videos of undergraduate lecture classes and make them available free to the public.”


Noted by JFL, MD
Statistical Literacy for Readers of Pediatrics: A Moving Target

Martha A. Hellems, MD, MSa, Matthew J. Gurka, PhDb, Gregory F. Hayden, MDa

Departments of aPediatrics and bPublic Health Sciences, University of Virginia, Charlottesville, Virginia

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

ABSTRACT

OBJECTIVE. Pediatric residents are expected to study research design and statistical methods to enable them to critically appraise the pediatric literature and apply the findings to patient care. However, it is not clear how best to teach these skills or even which statistical concepts are most important. An earlier study demonstrated that the statistical complexity of articles published in Pediatrics increased from 1952 to 1982. The goals of our study were to assess whether this trend has continued and to determine the statistical measures and procedures most commonly encountered in Pediatrics.

METHODS. We reviewed the print research articles published in Pediatrics, volume 115, 2005, and recorded the statistical measures and procedures reported in each article to determine how many articles used statistics or statistical procedures and what statistical procedures were encountered most commonly.

RESULTS. The proportion of articles that used any inferential statistics increased from 48% in 1982 to 89% in 2005. The mean number of inferential procedures per article increased from 2.5 in 1982 to 3.9 in 2005. The most commonly encountered statistical procedures or measures were descriptive statistics, tests of proportions, measures of risk, logistic regression, t-tests, nonparametric tests, analysis of variance, multiple linear regression, sample size and power calculation, and tests of correlation. However, a reader who is familiar with only these concepts can understand the analyses used in only 47% of articles.

CONCLUSIONS. Our results confirm a trend toward the use of new and increasingly complex statistical techniques in Pediatrics. Educational efforts might most profitably focus on the principles underlying statistical analysis rather than on specific statistical tests. Authors, reviewers, and journal editors have a greater responsibility for ensuring that statistical procedures are used appropriately, as it may be increasingly unrealistic to expect readers to fully understand the statistical analyses used in journal articles.
 Lifelong learning as a physician demands facility in the assessment and application of clinical evidence from the medical literature. Appraisal of an article’s methodologic rigor depends on an understanding of the study design and analysis used by the authors. The importance of these skills in medical training is broadly recognized. The Accreditation Council for Graduate Medical Education and the American Board of Pediatrics mandate that, to attain competency in practice-based learning and improvement, residents are expected to “apply knowledge of study designs and statistical methods to the appraisal of clinical studies” and to “appraise and assimilate evidence from scientific studies related to their patients’ health problems.”1,2 Optimally, graduates of pediatric residency programs will have a strong enough working knowledge of statistics to be able to evaluate the most important analyses in most of the medical literature they read. Unfortunately, graduates of pediatric residency programs nationwide often report that they receive little to no formal training in epidemiology and biostatistics, and they give only “fair to poor” marks for their knowledge of research design and statistical analysis.3

It is not clear how best to teach critical appraisal skills during pediatric residency training. Study design and biostatistics are often taught in the context of a journal club or an evidence-based medicine conference,4 but the success of such efforts in imparting these concepts has not been well evaluated.5 Time is limited, especially with new resident duty-hour restrictions, and the needs of the learners may vary widely. Medical students frequently have poor skills in basic mathematics, and they often have difficulty interpreting medical data.6 Residents, researchers, and practicing physicians may perform no better.7,8

It is also not clear which statistical concepts are most necessary and useful for readers to become familiar with. There has been a well-documented trend toward the use of new and increasingly complex statistical techniques in published articles.9 Use of these more sophisticated techniques can potentially allow more thorough analysis of study data by, for example, enabling complex modeling with multiple comparisons or multiple variables. But, such advances have made it more and more difficult for readers to understand the study analyses. An earlier study demonstrated that a reader of Pediatrics who understood descriptive statistics (for example, means and standard deviations) and 3 inferential statistical procedures (Student’s t test, χ², and Pearson’s r) could understand the statistical analysis in 97% of research articles published in 1952, but only 49% of articles in 1982.10 The goals of this study were to determine (1) whether this trend has continued and the proportion of articles that a reader can understand with only these few basic concepts has declined further, and (2) the statistical measures and procedures most commonly encountered in Pediatrics. These concepts could then potentially be used in planning a curriculum for pediatric residents and other readers wishing to improve their skills in critical appraisal of published research.

Methods
We reviewed the 171 print articles published in the Articles, Special Articles, and Review Articles sections of Pediatrics, volume 115 (January to June), 2005. To allow comparison with the results of the previous study, a single volume of Pediatrics was reviewed, exclusive of articles published online only. It was expected that exclusion of the online-only articles would not bias our results, because articles published online only are subject to the same peer-review process and selection criteria as the print Pediatrics.11 Volume 115 was chosen as the most recent complete volume at the time of initiation of our review, without any foreknowledge of the statistical content of the articles in this volume. Two of us (Drs Hellems and Hayden) jointly read the articles and recorded the statistical measures and procedures used in each article. Statistical measures and procedures were usually reported in the text of the methods and/or results sections, but were not infrequently found only in the tables, figures, or elsewhere. A statistician (Dr Gurka) helped to classify procedures with which the reviewers were unfamiliar. The tabulated measures and procedures were then categorized to determine (1) how many articles used statistics or statistical procedures, (2) how many articles had a statistical analysis that could be evaluated by a reader with an understanding only of descriptive statistics and 3 inferential statistical procedures (Student’s t test, χ², and Pearson’s r), and (3) which statistical procedures were encountered most frequently. The 3 inferential procedures chosen for the second analysis do not necessarily represent the most important statistics but were selected to allow comparisons with the previous study.

Results
A total of 171 articles were reviewed (Table 1). Only 1 article, a review published as a Special Article, used no statistics or statistical procedures. The proportion of articles that used only descriptive statistics declined from 23% in 1982 to 10% in 2005. Only 18% of articles in 2005 used no statistics or only descriptive statistics, Student’s t test, χ², and/or Pearson’s r, compared with 65% in 1982. If Review Articles and Special Articles are excluded from consideration, 9% of articles use only descriptive statistics, Student’s t test, χ², and/or Pearson’s r. The proportion of articles that used any inferential statistical procedure, with or without descriptive statistics, increased from 48% in 1982 to 89% in 2005. The mean number of inferential procedures per article increased from 2.5 in 1982 to 3.9 in 2005. Table 2 lists the statistical measures or procedures...
TABLE 1  Use of Statistical Procedures in Selected Volumes of Pediatrics, 1952–2005

<table>
<thead>
<tr>
<th>Statistical Measure or Procedure</th>
<th>No. (%) of Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of articles</td>
<td>171</td>
</tr>
<tr>
<td>No statistics, %</td>
<td>18 (10)</td>
</tr>
<tr>
<td>Only descriptive statistics, %</td>
<td>23 (13)</td>
</tr>
<tr>
<td>No statistics or only descriptive statistics, Student’s t test, χ², and/or Pearson’s r, %</td>
<td>89 (53)</td>
</tr>
<tr>
<td>Inferential statistical procedures, %</td>
<td>39 (23)</td>
</tr>
<tr>
<td>Mean inferential procedures/article</td>
<td>3.9 (2.4)</td>
</tr>
</tbody>
</table>


TABLE 2  Statistical Measures or Procedures Encountered in ≥10% of Articles Published in Pediatrics, Volume 115, 2005

<table>
<thead>
<tr>
<th>Statistical Measure or Procedure</th>
<th>No. (%) of Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive statistics (cited as mean, SD, variance, SE of the mean, percentage, confidence interval, median, percentile, range, histograms, t score, z score)</td>
<td>168 (98)</td>
</tr>
<tr>
<td>Tests of proportions (cited as Pearson χ², χ² test for trend, χ² test of homogeneity, McNemar’s test, paired χ², Fisher’s exact test, and unspecified tests of proportions)</td>
<td>92 (54)</td>
</tr>
<tr>
<td>Measures of risk (cited as odds ratio, relative risk, incidence ratio, likelihood ratio, mortality ratio, risk difference)</td>
<td>65 (38)</td>
</tr>
<tr>
<td>t test (cited as paired, unpaired, and 2-sided t tests, Satterthwaite method)</td>
<td>48 (28)</td>
</tr>
<tr>
<td>Logistic regression (includes 1 citation of “binormal regression”)</td>
<td>45 (26)</td>
</tr>
<tr>
<td>Nonparametric tests (cited as Wilcoxon rank sum, Mann-Whitney U, Mann-Whitney rank sum, Wilcoxon Mann-Whitney test for ordered categories, Wilcoxon signed-rank test, Kruskal-Wallis, Friedman’s statistic, median 2-sample test, median test for continuous data)</td>
<td>41 (24)</td>
</tr>
<tr>
<td>Multiple linear regression (includes simple linear regression, multiple linear regression, ANCOVA)</td>
<td>33 (19)</td>
</tr>
<tr>
<td>ANOVA (includes ANOVA, 2-way ANOVA, factorial ANOVA)</td>
<td>30 (18)</td>
</tr>
<tr>
<td>Sample size and power calculation</td>
<td>22 (13)</td>
</tr>
<tr>
<td>Tests of correlation (includes Pearson’s r, Spearman’s ρ, Spearman rank correlation)</td>
<td>19 (11)</td>
</tr>
</tbody>
</table>

ANOVA indicates analysis of variance; ANOVA, analysis of variance.

TABLE 3  Statistical Measures or Procedures Encountered in 5% to 9% of Articles Published in Pediatrics, Volume 115, 2005

<table>
<thead>
<tr>
<th>Statistical Measure or Procedure</th>
<th>No. (%) of Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repeated measures analysis</td>
<td>15 (9)</td>
</tr>
<tr>
<td>Transformation</td>
<td>14 (8)</td>
</tr>
<tr>
<td>Cluster analysis</td>
<td>12 (7)</td>
</tr>
<tr>
<td>Tests of agreement/reliability (includes κ, Cohen’s κ, Krippendorf’s κ, intraclass correlation coefficient)</td>
<td>12 (7)</td>
</tr>
<tr>
<td>Sensitivity/specificty (includes sensitivity, specificity, positive predictive value, negative predictive value, ROC curve)</td>
<td>11 (6)</td>
</tr>
<tr>
<td>Survival analysis (includes Cox regression, Kaplan-Meier, log-rank, hazard ratio, smoothed instantaneous hazard rate)</td>
<td>10 (6)</td>
</tr>
<tr>
<td>Cronbach’s α</td>
<td>9 (5)</td>
</tr>
<tr>
<td>Generalized estimating equations</td>
<td>9 (5)</td>
</tr>
<tr>
<td>Multiple comparison tests (includes Tukey-Kramer test, Tukey test, Scheffe’s test, Dunn’s method for nonparametric data, Bonferroni multiple comparison test)</td>
<td>8 (5)</td>
</tr>
<tr>
<td>Cost analysis</td>
<td>8 (5)</td>
</tr>
</tbody>
</table>

ROC indicates receiver operating characteristic.

encountered in ≥10% of the articles reviewed. A reader who understands all of these “top 10” topics can potentially understand the analyses used in only 47% of the 171 articles. Table 3 lists the procedures or measures that were encountered in 5% to 9% of articles. A reader who understands the concepts in Tables 2 and 3 can potentially understand the analyses used in only 70% of the 171 articles. Many other statistical techniques were encountered (Table 4).

Multivariable modeling techniques were encountered frequently. The most commonly used multivariable technique was logistic regression. Simple linear regression was not commonly used; in most instances, it was used to assess the impact of single independent variables before conducting the main multivariable analysis. When modeling was used, methods such as variable selection, imputation, model validation, and techniques for adjustment and standardization were often reported.

Nonparametric techniques were used in 24% of articles, most commonly the Mann-Whitney U test and the Kruskal-Wallis test. To add to potential confusion, multiple names were used for the Mann-Whitney U test, including Wilcoxon rank sum, Mann-Whitney U, Mann-Whitney rank sum, and Wilcoxon Mann-Whitney test for ordered categories.

Statistical methods were not always explained or even mentioned in the methods section of the articles, but were often buried in the text of the results section or listed only as footnotes to tables. In several instances, no statistical procedure was specified, but the presence of a p value indicated that a test had been performed. In most of these cases, it was possible to make an educated guess about what sort of procedure had been performed (eg, a test to compare proportions), but it was not possible to
determine which specific test had been used (eg, $\chi^2$ or Fisher’s exact test). In these cases, only the more general classification was tabulated.

**DISCUSSION**

This analysis of articles recently published in *Pediatrics* documents that the use and complexity of statistical analysis has increased over the past 50 years. The 1983 study concluded, “To understand the statistical aspects of current articles requires familiarity with a broader range of sophisticated statistical techniques than was necessary just a few years ago.” Our study documents that this trend has continued, and that the challenge of understanding the statistical analyses in published articles has only become greater. Virtually all articles now include some statistical measures or techniques, and most use inferential statistical procedures. The number of statistical measures and procedures used in articles has increased; a reader familiar only with the 10 most common statistical concepts will encounter unfamiliar techniques found in this 1 volume of *Pediatrics*. The breadth of techniques found in this 1 volume of *Pediatrics* is remarkable. (Parenthetically, each of this study’s authors, 1 of whom is a biostatistician, encountered tests with which he or she was unfamiliar, and an anonymous reviewer commented, “I have never heard of some of these and I teach this stuff!”) In addition, the types of analysis are changing. For example, although many biostatistics courses teach linear regression, logistic regression is much more commonly encountered in *Pediatrics*. This increasing complexity and “moving target” complicate efforts to determine a workable statistics curriculum for residents and practicing pediatricians.

Reasons for this increase in statistical complexity may include the development of new study designs and statistical techniques, and also the broad availability of expanded computing power. Perhaps this increased complexity of statistical analysis should be expected given the increasing complexity of the world in general, and of scientific domains in particular. Taken in this larger context, the statistical complexity is perhaps better understood, but it nevertheless may remain troublesome and baffling to readers.

A reader may understand a research article at several levels. He or she may understand the statistical tests and procedures well enough to assess whether they were appropriate to the study and conducted correctly, or he or she may be able to interpret results reported as descriptive statistics, measures of effect size, or $P$ values without understanding the statistical procedures used. The latter reader may still find an article to be valuable. If, however, one assumes that a general reader should be able to understand the statistical procedures and measures in most published articles, there are a few possible courses of action. One option might be for journals to require that statistical methods be kept relatively simple and that any unusual or complex procedures be explained thoroughly. In this context, “unusual” could be defined mathematically, for example, as a test appearing in <5% of articles. Such a requirement might, however, “dumb down” the techniques used, result in suboptimal analysis of study data, and increase the length of methods section that few would ever read, let alone comprehend.

The optimal way to report statistical methods no doubt depends on the article’s anticipated audience. Un-
fortunately there may be many audiences (or a continuum of audiences) based on readers’ levels of interest in the clinical topic and expertise in research design and analysis. For example, clinicians may have better understood the results of 1 reviewed study because it included helpful background information about the statistical model, as follows:

The Cox regression technique takes account of variable length of follow-up monitoring, including the possibility of “censoring” (no event when last observed but future events are not ruled out), and produces an estimate of the relative likelihood of the event during any small time interval (“hazard ratio”), as affected by specified risk factors. Like the conventional techniques of multiple linear and logistic regression, Cox regression can assess the independent effect of each risk factor while controlling simultaneously for other factors.13

This same information, however, may have been boring and superfluous for a reader with substantial statistical expertise. In contrast, statistically savvy readers may appreciate having substantial detail of a mathematical model, whereas most clinicians are unlikely to delve into a discussion of 6 different methods used to impute missing study data included in another reviewed study.14 Perhaps ideally, articles will include a brief overview of the statistical methods used, as well as significant detail (perhaps in an appendix) for statistical reviewers and any interested readers. In the instance of printed articles, additional information can be made available on request. For articles published electronically, readers who desire more information about the statistical technique or model could perhaps click on a link to access that material.

A second option would be to provide readers with more intensive training in statistical methods. Given current duty-hour restrictions for residents, however, finding more time to teach this material during residency will be difficult. Likewise, educational sessions on biostatistics at continuing medical education meetings are not likely to attract large audiences if they are competing with clinical updates or sessions on such practical issues as new vaccines or office management. Placing greater emphasis on teaching biostatistics to medical students is a possibility, but the practical value of this information may be less clear and, therefore, less interesting to students at this earlier stage of training.

A third option is to concede that many readers will never be motivated and/or able to understand the statistical analysis of most published articles. In past years when the variety of statistical techniques encountered was narrower than today, motivated physician readers could develop a rudimentary understanding of the techniques they were likely to encounter in published articles. Now that the range of techniques encountered has broadened so widely, the expectations may need to change. The purely statistical aspects of biomedical research are certainly not as important and as crucial to good science as is sound research design with attention to potential sources of bias, choice of appropriate controls, and types of outcomes chosen. Educational efforts focusing on principles of study design and potential biases might aid the clinician reader regardless of complexity of statistical analysis. A complementary approach is for clinicians to become “information masters” who efficiently use the medical literature, including secondary sources such as the Cochrane Database of Systematic Reviews, as well as assessments of the strength of research evidence, such as the strength of recommendation taxonomy.15,16 For most readers, understanding the “what” and “why” of the research is more important than understanding the “how” of the analysis.

Readers who do not understand the statistical measures and analysis used in an article have several options. Because ignorance often breeds mistrust, readers may tend to reject an unfamiliar analysis and discount an article’s results, but this might well result in dismissing an important research finding. Consulting a statistician for assistance may be helpful, but this is impractical for most readers. Reading an expert review of the article may be helpful, if one has been written. Trusting a study’s authors and the journal’s peer-review process to assure that the statistical analysis is appropriate and correct is another possibility, but journal editors may not conduct statistical reviews of submitted manuscripts,17 and statistical errors have been detected commonly in published articles.18–20

Including a biostatistician among the authors of an article probably increases the possibility that an “unfamiliar” statistical test is used, but may well also increase the likelihood that the analysis is thoughtful and appropriate.21 Including a statistician on editorial boards and having articles refereed by a statistician may make it “safer” for statistically naive readers to believe what they read.

This study has several limitations. First, only 1 volume of 1 journal was reviewed, and we excluded the electronic pages, thus the findings may not be generalizable to other journals. For example, a review published in 2003 of 6 journals in 3 nonpediatric subspecialties revealed that a reader could understand 70% of articles with 3 basic concepts: descriptive statistics, χ²/Fisher’s exact test, and Student’s t test.22 Pediatrics was reviewed for this study to allow comparisons with the earlier article.10 The study results may still be broadly applicable because, as the official journal of the American Academy of Pediatrics, Pediatrics has a large circulation and high impact factor, and publishes many articles of interest to both clinicians and researchers. A second limitation is that some statistical procedures actually used in the reviewed articles may have been missed in our review. In that case, our findings can only underestimate the frequency and complexity of statistical procedures that a reader might encounter. Third, no attempt was made to assess the appropriateness or accuracy of the statistical
measures and techniques used in each article. Finally, our classification of the statistical measure and procedures represents just 1 possible categorization. The concepts might be grouped in different ways.

CONCLUSIONS
This study demonstrates the increasing complexity of statistical analyses encountered in Pediatrics. Goals for education of pediatric residents and other readers may need to be reassessed to emphasize the understanding of principles of statistical inference rather than on the statistical procedures themselves. Authors, reviewers, and journal editors have a greater responsibility for ensuring that statistical procedures are used appropriately, because it may be increasingly unrealistic to expect readers to fully understand the statistical analyses used in journal articles.

ACKNOWLEDGMENT
We thank Dr Michael S. Kramer for thoughtful comments and suggestions.

REFERENCES
Relationship Between Angiotensin-Converting Enzyme Gene Insertion or Deletion Polymorphism and Insulin Sensitivity in Healthy Newborns

Tongyan Han, MD, PhD, Xinli Wang, MD, PhD, Yunpu Cui, MD, PhD, Hongmao Ye, MD, Xiaomei Tong, MD, Meihua Piao, MD

Department of Pediatrics, Third Hospital, Peking University, Beijing, China

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

ABSTRACT

CONTEXT. It was proposed that the association between low birth weight and adult insulin resistance is principally genetically mediated. The insertion/deletion polymorphism of the angiotensin-converting enzyme gene was associated with insulin sensitivity in adults.

OBJECTIVE. Our goal was to investigate the relationship between angiotensin-converting enzyme gene insertion/deletion polymorphism and the insulin sensitivity in healthy newborns.

PATIENTS AND METHODS. One hundred eighty healthy newborns, all of whom had a 1-minute Apgar score of >7 and gestational age >33 weeks, were enrolled in the study. Fasting glucose and insulin levels were measured on day 2 or 3 after birth, and angiotensin-converting enzyme genotype was determined.

RESULTS. The observed frequency distribution of angiotensin-converting enzyme genotypes did not deviate from that predicted by Hardy-Weinberg equilibrium in this group. There were no statistically significant differences in birth size and shape in different angiotensin-converting enzyme genotypes. Those carriers of the genotype homozygous for the deletion allele had the highest logarithmically transformed homeostasis model assessment compared with those who were heterozygous or homozygous for the insertion polymorphism. When compared with those with ≥1 insertion allele, those of the genotype homozygous for the deletion allele had significantly higher logarithmically transformed fasting insulin and logarithmically transformed homeostasis model assessment results. Regarding birth weight, birth length, ponderal index, and fasting glucose concentration, there were no significant differences between the genotype homozygous for the deletion allele and the genotypes heterozygous or homozygous for the insertion allele.

CONCLUSIONS. In this study, the deletion allele was associated with relatively impaired insulin sensitivity in healthy neonates. It may be a clue to explain the association between the deletion allele and insulin resistance in the long-term.
Impaired insulin sensitivity and insulin resistance are thought to contribute to the development of the pentad of hypertension, hyperinsulinism, dyslipidemia, obesity, and cardiovascular disease, known as metabolic syndrome.\(^1\) There is increasing evidence that insulin resistance is programmed during fetal development. In 1992, Hales and Barker\(^2\) proposed the “thrifty phenotype hypothesis,” which postulates that all features of metabolic syndrome have a strong environmental basis. It suggests that fetal and early nutrition play an important role in determining the susceptibility of an individual to these diseases. Since then, most studies about low birth weight infants proved that a poor intrauterine environment starts a process of adaptation (programming) to unfavorable environments in the fetus, and this process asserts itself especially at the hormonal level, such as in adrenal medulla, the hypothalamus-pituitary axis, and the renin-angiotensin systems (RASs). Seven years later, Hattersley and Tooke\(^3\) proposed that the association between low birth weight and adult insulin resistance was principally genetically mediated. Low birth weight, measures of insulin resistance in life, and ultimately glucose intolerance, diabetes, and hypertension would all be phenotypes of the same insulin resistant genotype. Both genetics and the fetal environment are likely to be important in determining fetal growth in the same way that both genetic and environmental influences are important in adult disease susceptibility.

RASs play an important role in circulatory homeostasis. Angiotensin-converting enzyme (ACE) generates angiotensin II, a pressor, directly through vasoconstriction and indirectly through stimulation of adrenal aldosterone release and resultant salt and water retention. In addition, ACE degrades vasodilator kinins. The insertion/deletion (I/D) polymorphism of the ACE gene is characterized by the presence (I) or absence (D) of a 287-base pair alu repeat sequence within intron 16 of the ACE gene. This polymorphism accounts for nearly half the variance in serum ACE levels.\(^4\) Tissue ACE activity is similar among those heterozygous for the I allele (ID) and homozygous for the I allele (II), with homozygous for the D allele (DD) associated with an elevation in tissue ACE activity of as much as 75% in the heart and white blood cells.\(^5,6\) The increased ACE level was suggested by some but not all studies to predispose to severe common cardiovascular and renal diseases,\(^7-9\) especially in people with diabetes.\(^9,10\) The view that the RAS is contained in the placenta and that the RAS is a factor in fetal growth as shown by the identification of receptors in trophoblast layers has gained significance.\(^11\) As in adult studies, an association between the RAS or DD carriers and insulin sensitivity was reported.\(^12,13\) Cambien et al\(^14\) found that the ACE genotypes were associated insulin response among a group of young adults who were born small-for-gestational-age, which supported the “fetal insulin hypothesis.”\(^1\) However, their results were confined to small-for-gestational-age infants. In our study, we explored whether there was an association between ACE gene I/D polymorphism and insulin sensitivity in healthy newborns to find some clues of genetic basis of insulin resistance that excluded the influences of environments (intrauterine, diet, lifestyle, and disease). Therefore, the aims of our study were to (1) measure fasting insulin (FI) and fasting glucose levels in healthy newborns, (2) determine the relation between ACE gene polymorphism and infant birth weight and shape, and (3) investigate the relation between ACE gene polymorphism and insulin sensitivity in healthy newborns.

**METHODS**

**Sample**

The subjects in our study were recruited from singleton newborns who were delivered from April through December 2003 in the department of obstetrics of the Third Hospital, Peking University. Newborns were included in this study if they fulfilled the following inclusion criteria: (1) they had experienced a normal pregnancy with gestational age of $>$33 weeks; (2) they had a 1-minute Apgar score of $>$7 and a 5-minute Apgar score of 10; (3) they were breastfed during the study; and (4) there was a genomic DNA sample that could be used for genotyping. Newborns were excluded if they were born to women with diabetes, gestational diabetes, gestational hypertension, or chronic hypertension and they had intrauterine infections and congenital malformations. All studies were performed after parents gave written informed consent; the study protocol was approved by the Third Hospital Ethics Committee. The investigation conformed to the principles outlined in the Declaration of Helsinki as revised in 2000.

**Measures**

**Birth Weight and Length**

Midwives measured the birth weights and crown-to-heel lengths within 2 hours of delivery. Birth weights were recorded to the nearest gram by using a balance scale.

**Fasting Glucose and Insulin Concentrations**

All of the neonates were breastfed since birth. Blood was obtained by heel prick before feeding between 7:00 and 9:30 AM on day 2 or 3 of life ($\geq$3 hours’ fast) and analyzed for glucose and insulin concentration.

Glucose concentrations were measured by using the SureStep Plus System from LifeScan (Milpitas, CA). Interassay and intraassay coefficients of variation for glucose were 0.9% and 1.8%, respectively. Insulin was measured by enzyme-amplified immunoassay using active insulin ELISA Kit (DSL-10-1600; Diagnostic Systems Laboratories, Webster, TX). The detection limit of this
assay was 0.26 μIU/mL (1.81 pmol/L) in our laboratory, and the intraassay and interassay coefficients of variation were 2.6% and 5.2%, respectively.

**Genotype of ACE Gene**

Genomic DNA was prepared from heel-prick blood (300 μL) by using a commercially available DNA isolation kit (Wizard genomic DNA purification kit; Promega, Madison, WI). The DNA concentration was adjusted to 100 ng/μL by adding distilled water. The presence of the insertion and deletion allele in intron 16 of the ACE gene was detected using the method of Rigat et al. The sequence of sense oligonucleotide primer was 5’-CTG GAG ACC ACT CCC ATC CTT TCT-3’ and the antisense primer 5’-GAT GTG GCC ATC ACA TTC GTC AGA-3’.

The polymerase chain reaction (PCR) was performed with 200 ng of genomic DNA template in a final volume of 25 μL containing 1.5 mmol/L MgCl₂, 50 mmol/L KCl, 10 mmol/L Tris-HCl, 10 pmol of each primer, 200 μmol/L of each deoxyribonucleotide triphosphate, and 1 unit of Taq DNA polymerase (Takara, Shiga, Japan). Amplification was performed by using a DNA thermal cycler (Gene Amp PCR System 9700; Perkin-Elmer, Foster City, CA) with 30 seconds denaturation at 94°C, 45 seconds annealing at 56°C, and 1-minute extension at 72°C for 35 cycles. In the last cycle, the extension step was conducted for 6 minutes. To avoid mistyping of heterozygotes as DD homozygotes, all DD genotype samples were confirmed by using a pair of primers that produce an amplified product only in the presence of the insertion, which was used to verify the polymorphism: sense 5’-TGG GAC CAC AGC GCC CGC CAC TAC-3’ and antisense 5’-TCG CCA GCC CTC CCA TGC CCA TAA-3’. The PCR condition was similar to that procedure for I/D detection, except that the annealing temperature was changed to 67°C. All PCR products were visualized after electrophoresis on a 1.5% agarose gel and ethidium bromide staining. Genotyping was performed in a blinded fashion.

**Statistical Analysis**

The previously validated homeostasis model assessment (HOMA) was used to estimate insulin sensitivity. HOMA was calculated from the fasting glucose and insulin concentrations according to the equation: HOMA = [insulin (μIU/mL) × glucose (mmol/L)]/22.5. Birth size and shape measures were birth weight, birth length, and ponderal index (PI = [birth weight/birth length³] × 100), studied as continuous variables.

The data are expressed as mean ± SD. All statistical analyses were performed by using the Statistical Package for Social Science 10.0 for Windows (SPSS Inc, Chicago, IL). Fasting insulin and HOMA were logarithmically transformed (log₁₀) before the analysis to approach normal distribution.

The subjects were primarily divided into 3 groups according to the 3 ACE genotypes. The statistical difference in genotype distribution and allele frequencies among the groups was assessed by the Pearson χ² test. One-way analysis of variance was used to test for differences in means of phenotypic characteristics between the 3 genotypes (with Bonferroni correction). We further combined the subjects of ID and II genotypes and compared with the DD carriers. The clinical characteristics of the 2 groups were compared by unpaired Student’s t test. P < .05 was considered statistically significant.

**RESULTS**

One hundred eighty newborns, including 135 term infants and 45 preterm infants, were taken into the scope of the study. All of them were Chinese and they were born at term or near term after a normal pregnancy. They had no major neonatal problems and had normal acid-base status at birth. There was no history of maternal hypertension, diabetes, or infections. Mean gestational age and birth weight of the study population were 37.65 ± 2.16 weeks and 2946.46 ± 645.75 g, respectively. The male/female ratio was 101:79.

All newborns were genotyped for the ACE I/D polymorphism. Because the ACE genotype distributions were not significantly different between term infants and preterm infants (χ² = 1.090; P = .380), we combined their data and analyzed them as 1 group. The frequencies of DD, ID, and II genotypes were 18.3%, 46.7%, and 35.0%, respectively. The allele frequencies were 41.67% and 58.33% for the D and I alleles, respectively. These results were consistent with the Hardy-Weinberg equilibrium (χ² = 0.188; P = .910). Demographic characteristics of gender, maternal age, gestational age, delivery method, and 1-minute Apgar score were not different between genotype groups (Table 1).

There were no statistically significant differences in birth size and shape in different ACE genotypes (Table 2). And the fasting glucose and fasting insulin (logarithmically transformed) were not significantly different among the 3 genotypes. The individuals with DD genotype had the highest HOMA (log₁₀HOMA = 0.21 ± 0.45) compared with individuals with the ID genotype (log₁₀HOMA = 0.01 ± 0.38; P = 0.016) or homozygous (II) (log₁₀HOMA = 0.01 ± 0.40; P = 0.021). After using Bonferroni correction, only the difference between DD genotype and ID genotype was significant.

As shown in Table 2, data for those of ID and II genotypes did not differ. When compared with those with ≥1 I allele, those with DD genotype had significant higher log₁₀FI (0.93 ± 0.41 vs 0.76 ± 0.36; P = .018) and log₁₀HOMA (0.21 ± 0.45 vs 0.01 ± 0.38; P = .010) (Fig 1). Regarding birth weight, birth length, PI, and fasting glucose concentration, there were no significant differences between DD genotype and ID+II genotypes.
Recruits to this study were born at term or near term after a normal pregnancy, had normal acid-base status at birth, no history of maternal hypertension, and were breastfed during the study. The data were, therefore, not confounded by factors that were reported previously to be associated with serum insulin. The frequencies of ACE genotypes and the frequency of D allele in the study population were not different from those reported in our local population. Despite this relatively homogeneous study population, an evident increased fasting insulin and HOMA (logarithmically transformed) in the DD genotype infants was noted when compared with ID+II genotypes, which suggested an association between the D allele and relatively impaired insulin sensitivity.

The results of our study raise the possibility that the ACE genotype is related to insulin sensitivity. However, previous adult studies have reported conflicting results. Perticone et al\(^1\) reported that in a group of never-treated hypertensive individuals with the DD genotype were more insulin resistant as determined by the HOMA method than those individuals with either the ID or II genotype groups. Katsuya et al\(^2\) reported that normal subjects with the DD genotype were more insulin sensitive and had a lower insulin response to oral glucose administration. Panahloo et al\(^3\) found no influence of ACE genotypes on insulin sensitivity evaluated by plasma proinsulin levels and HOMA in nondiabetic subjects.

The ability of the ACE genotype to influence glucose metabolism is not understood; 1 possible mechanism explaining the link may be the elevated ACE levels that are associated with the DD genotype. Genetic analyses on I/D polymorphism of the ACE gene showed that circulating and tissue ACE levels were higher in subjects with the DD genotype than in subjects with other genotypes.\(^4\) However, because the ACE I/D polymorphism is intronic, the mechanism of ACE overexpression in subjects with DD genotype is unclear; it is possible that this relationship is the result of tight linkage to another locus involved in the regulation of ACE gene expression.\(^5\)

Previous studies found that elevated ACE levels were associated with diabetes mellitus.\(^6\) As the main metab-

### TABLE 1  Comparison of Demographic Characteristics According to ACE Genotypes

<table>
<thead>
<tr>
<th></th>
<th>ACE Genotypes</th>
<th>(\chi^2/F)</th>
<th>(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DD (n = 33)</td>
<td>ID (n = 84)</td>
<td>II (n = 63)</td>
</tr>
<tr>
<td>Male, n (%)(^a)</td>
<td>21 (63.6)</td>
<td>40 (47.6)</td>
<td>40 (63.5)</td>
</tr>
<tr>
<td>Maternal age, mean ± SD, y(^b)</td>
<td>28.79 ± 5.04</td>
<td>28.49 ± 5.05</td>
<td>28.84 ± 4.04</td>
</tr>
<tr>
<td>Vaginal delivery, n (%)(^a)</td>
<td>22 (67.3)</td>
<td>62 (73.8)</td>
<td>44 (69.8)</td>
</tr>
<tr>
<td>Gestational age, mean ± SD, wk(^c)</td>
<td>37.51 ± 2.08</td>
<td>37.96 ± 2.15</td>
<td>37.31 ± 2.18</td>
</tr>
<tr>
<td>Apgar score at 1 min, mean ± SD(^d)</td>
<td>9.73 ± 0.63</td>
<td>9.74 ± 0.64</td>
<td>9.78 ± 0.58</td>
</tr>
</tbody>
</table>

Comparisons were performed with \(^a\) Pearson’s \(\chi^2\) or \(^b\) analysis of variance.

### TABLE 2  Comparison of Birth Size and Shape and Metabolic Characteristics According to ACE Genotypes

<table>
<thead>
<tr>
<th></th>
<th>ACE Genotypes</th>
<th>(F)</th>
<th>(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DD (n = 33)</td>
<td>ID (n = 84)</td>
<td>II (n = 63)</td>
</tr>
<tr>
<td>Birth weight, g</td>
<td>2858.18 ± 616.70</td>
<td>2982.50 ± 638.95</td>
<td>2944.63 ± 674.64</td>
</tr>
<tr>
<td>Birth length, cm</td>
<td>48.09 ± 3.46</td>
<td>48.67 ± 2.77</td>
<td>48.29 ± 3.10</td>
</tr>
<tr>
<td>PI</td>
<td>2.54 ± 0.26</td>
<td>2.55 ± 0.29</td>
<td>2.57 ± 0.25</td>
</tr>
<tr>
<td>Fasting glucose, mmol/L</td>
<td>4.44 ± 1.40</td>
<td>4.14 ± 0.94</td>
<td>4.10 ± 1.08</td>
</tr>
<tr>
<td>FL, mIU/L</td>
<td>13.22 ± 14.52</td>
<td>7.99 ± 8.09</td>
<td>7.85 ± 5.10</td>
</tr>
<tr>
<td>(\log_{10})FI</td>
<td>0.93 ± 0.41</td>
<td>0.75 ± 0.37</td>
<td>0.77 ± 0.38</td>
</tr>
<tr>
<td>HOMA</td>
<td>3.03 ± 5.34</td>
<td>1.46 ± 1.46</td>
<td>1.46 ± 1.05</td>
</tr>
<tr>
<td>(\log_{10})HOMA</td>
<td>0.21 ± 0.45</td>
<td>0.01 ± 0.38</td>
<td>0.01 ± 0.40</td>
</tr>
</tbody>
</table>

Data are expressed as mean ± SD. Fasting insulin and HOMA were logarithmically transformed (\(\log_{10}\)) before the analysis to approach normal distribution. All statistical comparisons were performed with analysis of variance. — indicates not applicable.
olle of serum ACE activity, angiotensin II was found to be a modulator of insulin sensitivity in both diabetic and nondiabetic subjects.24,25 ACE inhibitors improve insulin sensitivity in noninsulin dependent diabetes mellitus patients,26 nondiabetic hypertensive patients,27 and nondiabetic normotensive subjects.28 These findings suggest the possibility that elevated ACE levels are associated with reduced insulin sensitivity and glucose intolerance, which contributes to the development of metabolic syndrome.

It was proposed that the RAS has a pivotal role in fetal development and growth, which is contained in the placenta as shown by the identification of receptors in trophoblast layers.11,29 Studies have shown that the RAS and angiotensin II are fetal growth factors that have a significant role in the regulation of uteroplacental blood flow by means of angiotensin receptors, as well as in decidualization, placentation, and implantation.31,29 Moreover, ACE activation was suggested to result in redistribution of the placental circulation and thus probably reduced nutrient transfer to the fetus.30 Therefore, the carrier of DD genotype, with relatively higher ACE concentration, is associated with increased risk of growth restriction. In our study, we failed to find any significant differences in birth size and shape between 3 ACE genotypes, although the carriers of DD genotype had a tendency to be lighter and smaller. This may because of the limited number of subjects examined and only healthy newborns were included in the study.

CONCLUSIONS

In this group of healthy term and near-term neonates, DD genotype carriers had a significantly increased fasting insulin and HOMA, which suggested that the D allele was associated with relatively impaired insulin sensitivity. This fact may provide genetic evidence for the clustering of metabolic syndrome or insulin resistance syndrome and may be a clue to explain the association between the D allele and insulin resistance in the long-term. Because of the limitation of the study, we believe that a larger sample size with more small-for-gestation-al-age infants as a control group addressing the impact of ACE genotype on insulin sensitivity could provide a more robust result.

REFERENCES


READERS ARE CHANGING

“The news about newspapers could hardly be more dismal: falling circulation, repeated rounds of layoffs, disappearing ads and a chain of bad earning reports. It’s an unsavory stew of ills, one that shows little prospect of becoming more appetizing. . . . Perhaps most worrisome is the loss of young readers, who have drifted away steadily since the early 1970s, long before there was an Internet, when more than 70% of 18- to 34-year-old Americans read a daily newspaper. Last year that figure stood at 35%. . . . The time that Americans spend reading newspapers has been dropping steadily (now down to 15 hours a month), with scant evidence that quality Internet time is taking its place. In September, the average visitor to newspaper Web sites spent only 41.5 minutes per month on those sites, up 10% from the previous year but not nearly enough to make up the loss. . . . And while the use of newspaper Web sites is growing, the vast preponderance of Americans get their online news through the big portals (AOL, Yahoo, etc.), which means that they are mostly consuming a bland porridge of wire service stories. . . . Most fundamental is whether the public is still interested in news (as opposed to entertainment, gossip or lifestyle info). More than fearing the death of newspapers—they will struggle on—we ought to fear what changing reading and viewing habits are forcing newspapers to think of as news. We shouldn’t fault the papers for this, however, any more than we should fault the evening news for going soft or the newweeklies for their endless lifestyle covers or CNN for its hyperventilating over every weather blip. They’re merely providing what their customers are demanding.”

Rattner S. *Wall Street Journal.* February 16, 2007
Noted by JFL, MD
ARTICLE

Infant Swimming Practice, Pulmonary Epithelium Integrity, and the Risk of Allergic and Respiratory Diseases Later in Childhood

Alfred Bernard, PhD, Sylviane Carbonnelle, MD, Xavier Dumont, BSc, Marc Nickmilder, PhD

Unit of Toxicology, Department of Public Health, Faculty of Medicine, Catholic University of Louvain, Brussels, Belgium

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

ABSTRACT

OBJECTIVE. Irritant gases and aerosols contaminating the air of indoor swimming pools can affect the lung epithelium and increase asthma risk in children. We evaluated the impact of infant swimming practice on allergic status and respiratory health later in childhood.

METHODS. Clara cell protein, surfactant-associated protein D, and total and aeroallergen-specific immunoglobulin E were measured in the serum of 341 schoolchildren aged 10 to 13 years, among whom 43 had followed an infant swimming program. Asthma was defined as doctor-diagnosed asthma and/or positive exercise-induced bronchoconstriction (15% decrease in postexercise forced expiratory volume).

RESULTS. There were no significant differences between the infant swimming group and the other children regarding the levels of exhaled nitric oxide and total or aeroallergen-specific serum immunoglobulin E. Children who swam as infants showed, by contrast, a significant decrease of serum Clara cell protein and of the serum Clara cell protein/surfactant-associated protein D ratio integrating Clara cell damage and permeability changes of the lung epithelial barrier. These effects were associated with higher risks of asthma and of recurrent bronchitis. Passive exposure to tobacco alone had no effect on these outcomes but seemed to interact with infant swimming practice to increase the risk of asthma or of recurrent bronchitis.

CONCLUSIONS. Our data suggest that infant swimming practice in chlorinated indoor swimming pools is associated with airways changes that, along with other factors, seem to predispose children to the development of asthma and recurrent bronchitis.
PUBLIC SWIMMING POOLS need to be disinfected, and usually this is done by water chlorination using sodium or calcium hypochlorite, chlorine gas, or chloroisocyanurates. These chlorine-based disinfectants, loosely referred to as “chlorine,” are added in water to release hypochlorous acid, a powerful oxidant that is the active biocide. When reacting with nitrogenous compounds originating from sweat, saliva, or urine brought by swimmers, hypochlorous acid generates a complex mixture of harmful byproducts, among which the most irritant are the chloramines. Monochloramine and dichloramine (NH₂Cl and NHCl₂, respectively) are water soluble, and the sum of their concentrations in water are referred to as combined chlorine. By contrast, the trichloramine also called nitrogen trichloride (NCl₃) is a water-insoluble gas that, once formed, is immediately released in the air, which gives indoor pools their distinctive “chlorine” smell. Depending on the bather’s hygiene and the pool ventilation, mean levels of trichloramine in community indoor pools in Europe are in the range of 300 and 800 μg/m³, which makes this gas one of the most concentrated air pollutants to which children of developed countries are regularly exposed. The air of swimming pools, particularly just above the water’s surface, is also contaminated by mist or aerosols laden with hypochlorous acid, combined chlorine, and other water-soluble chlorination byproducts, all increasing the burden of oxidants actively inhaled by swimmers.

The acute toxicity of chlorine-based disinfectants has been known for a long time. Case reports regularly describe lung injuries after community accidents in indoor swimming pools. Inhalation of chlorine gas is usually responsible for the acute lung damage that fortunately is usually transient with recovery of the lung function within a period of a few weeks. The possibility that the gaseous and aerosolized chlorination products building up in pool air can cause chronic respiratory effects in swimmers has, however, been acknowledged only recently. Our investigations on children attending indoor chlorinated swimming pools have shown that trichloramine, together probably aerosolized hypochlorous acid and chloramines, can damage the lung epithelium and promote the development of asthma, especially in atopic children. These effects might lie behind the strong ecological associations that we have recently evidenced across Europe between childhood asthma prevalence and the availability of indoor chlorinated swimming pool. Studies by other researchers confirmed the detrimental effects of these chemicals on the airways of recreational swimmers, including asthmatics, while providing additional evidence that swimming pool attendance during infancy might contribute to the development of allergic diseases.

One of the most critical factors in determining the risks of chlorination products for children seems to be the timing of exposure. The risk of developing asthma or of lung inflammation as assessed on the basis of exhaled nitric oxide (eNO) seems to culminate when children regularly attend indoor pools before the age of 7 years. A likely explanation for this higher sensitivity of young children is that they cannot really swim before the age of 6 or 7 years and, therefore, have to attend the small heavily polluted pool. When playing or learning to swim, young children probably also inhale and swallow more aerosols and water droplets containing hypochlorous acid and soluble chloramines. Another likely explanation is that the lungs of very young children are still developing, thus they are presumably more vulnerable to the irritating effects of chlorine and its derivatives.

These findings unavoidably raise the question of the safety of infant swimming, especially because the higher water temperature and the greater organic pollution in swimming pools attended by young children are conditions favoring the formation of chlorination byproducts. In the United States, it is estimated that between 5 and 10 million infants and preschool children participate in formal swimming instructions programs, and among them there are probably several million individuals who have learned swimming as infants. Surprisingly, most industrialized countries have popularized this practice in the absence of reliable data concerning the possible consequences of exposing infants to the toxic gases and aerosols building up in the air of indoor pools. The only study having addressed this issue is that of Nystad et al, who found that infant swimming practice was associated with an increased risk of recurrent respiratory tract infections and otitis media in the first year of life. This study, however, did not report data on the levels of chlorine in the pools attended by the infants, nor did it assess the effects of infant swimming on the development of asthma and allergy later in childhood.

In this cross-sectional study, we compared the respiratory health, allergic status, and pulmonary epithelium integrity of school-aged children who took part in an infant swimming program with that of their peers who did not. Whenever possible, objective outcome measures were used in addition to the traditional indicators of allergic and respiratory diseases assessed by questionnaire.

MATERIALS AND METHODS

Forty-three children having taken part in an infant swimming program were identified in a survey that involved a total of 341 schoolchildren 10 to 13 years of age. These children were recruited in 10 primary schools in southwestern Brussels. Levels of active and combined chlorine in the public swimming pool attended by swimming infants were within recommended limits at that time (<1.5 and 2 mg/L, respectively). Concentrations of trichloramine in pool air, available from 2001 when the systematic survey of this gas started in Brussels, ranged from 170 to 540 μg/m³ (mean: 325 μg/m³; n = 7).
Levels of trichloramine were probably in the same range some 10 years ago because operating conditions and recommended limits of chlorine had remained unchanged since the 1980s. All children were examined in their school between March 28 and May 29, 2002, thus outside main periods of pollination in Belgium. The protocol for examining children was described in detail elsewhere. Briefly, after their parents had given written informed consent, the children underwent a medical examination that included measurement of height and weight, collection of 1 blood sample (7 mL) after application of an anesthetic cream (EMLA, AstraZeneca, Karlskargo, Sweden). Asthma was screened by using the exercise-induced bronchoconstriction (EIB) test (on the basis of a 15% decrease of forced expiratory volume [FEV1] after a 6-minute roundabout indoors with submaximal effort), a noninvasive test that has been found to be predictive of clinical asthma or asthma symptoms in several studies. The examination also included the measurement of eNO and serum Clara cell protein (CC16) and surfactant-associated protein D (SP-D) markers of the deep lung epithelium integrity. Total immunoglobulin E (IgE) and IgE against the 12 most common aeroallergens were also measured in serum (Immulite Total and AlatTOP; Diagnostic Products Corp, Los Angeles, CA). Information about the respiratory health of children (respiratory symptoms during the previous 12 months, doctor-diagnosed asthma, and recurrent bronchitis) and their exposure to risk factors of asthma and allergy was obtained from the questionnaire filled in by the parents. The ethics committee of the Catholic University of Louvain approved the study protocol. Statistical comparison was based on the χ2 test, the 2-sided unpaired t test, or the Mann-Whitney U test. Bonferroni’s correction was applied to multiple comparisons. Backyard multiple and logistic regression analyses were used to assess associations between infant swimming practice and outcomes. We tested a total of 23 other potential predictors, including among others gender, total and aeroallergen-specific IgE, family history of allergic diseases, maternal smoking during pregnancy, parental smoking at home, siblings, housing density, exposure to pets, cumulated attendance at an indoor chlorinated pool, and the accessibility to a backyard chlorinated pool. In multiple regression analyses, cumulated pool attendance and the CC16/SP-D ratio that were not normally distributed were normalized by logarithmic transformation. To enable the log transformation, we set the minimum cumulated pool attendance at 2.5 hours instead of 0 (2.5 hours was the lowest reported cumulated pool attendance). Independent variables in multiple regression analyses were entered at a P value of .25 and kept in the model at P value <.05. In logistic regression analyses, adjustment was made for variables remaining with a P value <.1. These logistic regression analyses were repeated on children who were exposed to tobacco smoke during pregnancy or at home to detect the possible interactions of these 2 factors with infant swimming practice. Unless otherwise stated, only odds ratios (ORs) adjusted for covariates are reported. In addition, we used 2-way analysis of variance to assess the changes in serum pneumoproteins associated with infant swimming practice, access to a backyard pool, or with exposure to tobacco smoke during pregnancy or at home, and the possible interactions between these factors. Statistical analyses were performed by using SAS 9.1.3 (SAS Institute, Inc, Cary, NC).

RESULTS

As shown in Table 1, children who had been swimming as infants did not differ from their peers with respect to age, gender, ethnicity, BMI, and family history of asthma or hay fever. Children in the swimming infant group were also not significantly different from their peers regarding birth weight, number of siblings, housing density, or proportions of children who were breastfed, attended day care, or lived with pets since birth. The proportions of children exposed to tobacco smoke at home or who had been exposed to tobacco smoke during pregnancy, as well as of children having access to a backyard chlorinated pool were, however, noticeably greater among swimming infant children. These children also had a significantly greater cumulated attendance at indoor chlorinated swimming pools.

The allergic status and the markers of lung inflammation and epithelial integrity of children who swim as infants and controls are compared in Table 2. Mean levels of total serum IgE and the prevalences of IgE specific to major aeroallergens were similar between the 2 groups. The rate of sensitization to dog-specific IgE was higher in children in the infant swimming group, a difference, however, that was no more statistically significant after the application of the Bonferroni’s test. There were also no differences between the 2 groups regarding the mean levels and prevalences of elevated values of eNO. Most children positive in the eNO test (n = 29) were sensitized against house-dust mite allergen (n = 21 [72.9%]) and aeroallergen-specific IgE (n = 26 [89.7%]). Children in the infant swimming group, by contrast, showed a significant decrease of serum CC16 and an even more significant decrease of the CC16/SP-D ratio, an index integrating the damage to Clara cells and the permeability changes of lung epithelial barrier. In multivariate analysis, infant swimming emerged as the only statistically significant predictor of serum CC16 (partial r: −0.14; P = .01), and this practice was the strongest determinant of the CC16/SP-D ratio (log-transformed values, partial r: −0.15; P = .006), the latter being also influenced by day care attendance (partial r: −0.086; P = .007), age (partial r: −0.058; P = .009), breastfeeding (partial r: −0.080; P = .022), and
ethnicity (partial $r$: 0.087; $P = .024$). Passive exposure to tobacco smoke during pregnancy or at home as well as the access to a backyard pool had thus no influence on the serum levels of CC16 or on the serum CC16/SP-D ratio. Analysis of the effects of infant swimming and passive smoking on pneumoproteins by 2-way analysis of variance confirmed the decrease of serum CC16 and CC16/SP-D ratio in children in the infant swimming group and found no interaction between infant swimming and passive smoking (all $P > .15$).

As shown in Table 3, changes observed in serum pneumoproteins were associated with poorer respiratory health. Children who swam as infants showed an increased risk of chest tightness. There were no statistically significant differences in the other respiratory symptoms, which tended to be more prevalent in the children in the infant swimming group. These children were also ∼3 times more likely to be positive in the EIB test, to have doctor diagnosed and/or EIB test-screened asthma, and to suffer from recurrent bronchitis. Of note, the lack of statistical significance in the increased risk of doctor-diagnosed asthma with infant swimming was because of the influence of backyard pool (OR: 4.27; 95% confidence interval [CI]: 1.05–17.4). Indeed, removal of the backyard pool factor from the list of predictors increased the OR for doctor-diagnosed asthma associated with infant swimming to a level that was significantly $>1$ (OR: 2.96; 95% CI: 1.08–8.11). No significant association emerged between any of these outcomes or passive exposure to tobacco smoke at home or during pregnancy.

### TABLE 1  Characteristics of Children Who Swam as Infants and Their Controls

<table>
<thead>
<tr>
<th></th>
<th>Swimming Infants ($N = 43$)</th>
<th>Other Children ($N = 298$)</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD), ya</td>
<td>11.5 (0.6)</td>
<td>11.5 (0.6)</td>
<td>.98</td>
</tr>
<tr>
<td>Boys, n (%)</td>
<td>22 (51.1)</td>
<td>150 (50.3)</td>
<td>.92</td>
</tr>
<tr>
<td>White, n (%)</td>
<td>35 (81.4)</td>
<td>228 (76.5)</td>
<td>.48</td>
</tr>
<tr>
<td>BWL, mean (SD), kg/m^2</td>
<td>18.5 (2.7)</td>
<td>18.4 (2.5)</td>
<td>.81</td>
</tr>
<tr>
<td>Mother and/or father with asthma, n (%)</td>
<td>6 (14.0)</td>
<td>58 (19.5)</td>
<td>.39</td>
</tr>
<tr>
<td>Mother and/or father with hay fever, n (%)</td>
<td>13 (30.2)</td>
<td>87 (29.2)</td>
<td>.89</td>
</tr>
<tr>
<td>Birth weight, mean (SD), kg</td>
<td>3.19 (0.57)</td>
<td>3.34 (0.54)</td>
<td>.96</td>
</tr>
<tr>
<td>No. of siblings, mean (SD)</td>
<td>2.56 (1.30)</td>
<td>2.68 (1.28)</td>
<td>.56</td>
</tr>
<tr>
<td>Housing density, mean (SD), persons per room</td>
<td>0.83 (0.36)</td>
<td>0.79 (0.30)</td>
<td>.36</td>
</tr>
<tr>
<td>Breastfeeding, n (%)</td>
<td>9 (20.9)</td>
<td>34 (11.4)</td>
<td>.21</td>
</tr>
<tr>
<td>Exposure to pets since birth, n (%)</td>
<td>9 (20.9)</td>
<td>47 (15.8)</td>
<td>.39</td>
</tr>
<tr>
<td>Child care attendance n (%)</td>
<td>24 (55.8)</td>
<td>151 (50.7)</td>
<td>.40</td>
</tr>
<tr>
<td>Passive smoking at home, n (%)</td>
<td>18 (41.8)</td>
<td>72 (24.2)</td>
<td>.014</td>
</tr>
<tr>
<td>Maternal smoking during pregnancy, n (%)</td>
<td>16 (37.2)</td>
<td>46 (15.4)</td>
<td>.0005</td>
</tr>
<tr>
<td>Backyard pool, n (%)</td>
<td>6 (14.0)</td>
<td>9 (3.0)</td>
<td>.0011</td>
</tr>
<tr>
<td>Cumulated indoor pool attendance, median</td>
<td>146 (88–281)</td>
<td>83 (42–182)</td>
<td>.0002</td>
</tr>
<tr>
<td>(interquartile range), h</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 2  Total and Aeroallergen-Specific Serum IgE, eNO, and Serum Pneumoproteins in Children Who Swam as Infants and Their Controls

<table>
<thead>
<tr>
<th></th>
<th>Swimming Infants ($N = 43$)</th>
<th>Other Children ($N = 298$)</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total IgE, median (interquartile range), kIU/L</td>
<td>54.7 (24.6–162)</td>
<td>55.8 (21.9–175)</td>
<td>.96</td>
</tr>
<tr>
<td>Aeroallergen-specific IgE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panel of 12 aeroallergens, n(%)</td>
<td>13 (30.2)</td>
<td>95 (31.9)</td>
<td>.83</td>
</tr>
<tr>
<td>House dust mite, n(%)</td>
<td>6 (14)</td>
<td>57 (19.1)</td>
<td>.40</td>
</tr>
<tr>
<td>Cat, n (%)</td>
<td>2 (4.7)</td>
<td>17 (5.7)</td>
<td>.77</td>
</tr>
<tr>
<td>Dog, n (%)</td>
<td>4 (9.3)</td>
<td>8 (2.7)</td>
<td>.03</td>
</tr>
<tr>
<td>Pollen, n (%)</td>
<td>7 (16.3)</td>
<td>31 (10.4)</td>
<td>.26</td>
</tr>
<tr>
<td>eNO Median (interquartile range), ppb</td>
<td>10.2 (7.15–14.1)</td>
<td>8.9 (6.9–13.4)</td>
<td>.51</td>
</tr>
<tr>
<td>$&gt;30$ ppb, n(%)</td>
<td>3 (7.0)</td>
<td>26 (8.7)</td>
<td>.70</td>
</tr>
<tr>
<td>Pneumoproteins</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC16, mean (SD), $\mu$g/L</td>
<td>8.0 (3.3)</td>
<td>10.4 (4.2)</td>
<td>.01</td>
</tr>
<tr>
<td>SP-D, mean (SD), $\mu$g/L</td>
<td>113 (42)</td>
<td>100 (45)</td>
<td>.08</td>
</tr>
<tr>
<td>CC16/SP-D ratio, median (interquartile range)</td>
<td>0.07 (0.05–0.12)</td>
<td>0.10 (0.07–0.16)</td>
<td>.003</td>
</tr>
</tbody>
</table>

$a$ By 2-sided Mann-Whitney U test.
$b$ By $\chi^2$ test.
$c$ By 2-sided unpaired $t$ test.
$^d$ Statistically significant after application of the Bonferroni’s correction to the multiple comparisons of the lung epithelium markers (critical $P = .05/3$).
TABLE 3  Swimming Infant Practice, Respiratory Symptoms, and Risks of Asthma and Recurrent Bronchitis Later During Childhood

<table>
<thead>
<tr>
<th></th>
<th>Swimming Infants (N = 43)</th>
<th>Other Children (N = 298)</th>
<th>Crude (95% CI)</th>
<th>Adjusted (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheezing</td>
<td>6 (14.0)</td>
<td>26 (8.7)</td>
<td>1.7 (0.7–4.4)</td>
<td>1.5 (0.5–4.1)</td>
<td>.48</td>
</tr>
<tr>
<td>Chest tightness</td>
<td>6 (14.0)</td>
<td>13 (4.4)</td>
<td>3.6 (1.3–9.9)</td>
<td>3.8 (1.2–12.1)</td>
<td>.03</td>
</tr>
<tr>
<td>Shortness of breath</td>
<td>6 (14.0)</td>
<td>18 (6.0)</td>
<td>2.5 (0.9–6.8)</td>
<td>2.4 (0.8–7.3)</td>
<td>.11</td>
</tr>
<tr>
<td>Cough</td>
<td>11 (25.6)</td>
<td>51 (17.1)</td>
<td>1.7 (0.8–3.5)</td>
<td>1.8 (0.8–3.8)</td>
<td>.13</td>
</tr>
<tr>
<td>Doctor-diagnosed asthma</td>
<td>7 (16.3)</td>
<td>23 (7.7)</td>
<td>2.3 (0.9–5.8)</td>
<td>2.2 (0.7–7.6)</td>
<td>.10</td>
</tr>
<tr>
<td>Exercise-induced asthma (EIB)</td>
<td>4 (9.3)</td>
<td>11 (3.7)</td>
<td>2.7 (0.8–8.8)</td>
<td>4.3 (1.1–17.2)</td>
<td>.049</td>
</tr>
<tr>
<td>Doctor-diagnosed asthma and/or EIB</td>
<td>10 (23.3)</td>
<td>33 (11.1)</td>
<td>2.4 (1.1–5.4)</td>
<td>3.0 (1.3–7.4)</td>
<td>.01</td>
</tr>
<tr>
<td>Recurrent bronchitis</td>
<td>26 (60.5)</td>
<td>110 (36.9)</td>
<td>2.6 (1.4–5.8)</td>
<td>2.6 (1.3–5.1)</td>
<td>.006</td>
</tr>
</tbody>
</table>

a Adjusted for age, gender, atopy (IgE > 100 IU/mL), maternal smoking during pregnancy, and parental hay fever.
b Adjusted for BMI, sibling status, parental asthma, mold on the wall of room, and sport practice other than swimming.
c Adjusted for BMI, sibling status, maternal smoking during pregnancy, and parental asthma.
d Adjusted for breastfeeding.
e Adjusted for gender, parental asthma, aeroallergen-specific IgE, exposure to pets, number of siblings, and backyard pool.
f Adjusted for age, parental asthma, parental smoking at home and aeroallergens specific IgE.
g Adjusted for age, parental asthma; and aeroallergens specific IgE in serum.
h Adjusted for birth weight, day care attendance, house with double-gazed windows, and use of chlorine bleach for house cleaning.

except for the risk of wheezing that was increased by maternal smoking during pregnancy (OR: 2.75; 95% CI: 1.04–6.61). Passive exposure to tobacco smoke and infant swimming seemed, however, to potentiate the detrimental effects of infant swimming practice. The OR for asthma (doctor-diagnosed and/or screened with the EIB test) associated with infant swimming was indeed >2 times greater when children were also exposed to parental smoking (OR: 6.54; 95% CI: 1.18–36.4 vs OR: 2.52; 95% CI: 0.86–7.39) although the odds for recurrent bronchitis associated with infant swimming was >4 times greater among children who had been exposed to maternal smoking during pregnancy (OR: 6.99; 95% CI: 1.60–30.5 vs OR: 1.61; 95% CI: 0.71–3.65).

To determine whether changes in serum pneumoproteins were primarily because of the infant swimming practice and not the mere consequences of asthma or recurrent bronchitis that were more frequent in the children in the swimming infant group, we assessed by a 2-way analysis of variance the effects of infant swimming practice and not the mere consequences of asthma or recurrent bronchitis on serum levels of pneumoproteins, and the possible interactions between these factors. As illustrated in Fig 1, asthma and recurrent bronchitis alone had no influence on the serum levels of pneumoproteins, which is in sharp contrast with the highly significant decrease of serum CC16 and CC16/SP-D ratio associated with infant swimming activity (P < .001). This analysis clearly shows that the levels of serum CC16 were on average the lowest among the children in the infant swimming group who were asthmatic or recurrently suffered from bronchitis.

**DISCUSSION**

Our findings show that the infant swimming practice is associated with lung epithelium alterations that seem to predispose children to the development of asthma and recurrent bronchitis. These effects cannot be explained by inherited differences in the genetic disposition to develop these diseases, because children who swim as infants were well matched with the others with respect to the family history of respiratory disease, as well as to the serum levels of total IgE. Our findings cannot be explained either by the fact that on average the children in the infant swimming group were more exposed to parental smoking or to chlorine from backyard pools. In our study, we found no evidence that infant swimming increased the risk of respiratory allergy. There was also no increase in the risk of airways inflammation as assessed by the eNO test, a finding that is not surprising given the very close relationship between eNO and allergic sensitization observed by us and also reported by recent studies.28,29

As with any epidemiologic study using data from self-administered questionnaires, our study can be prone to recall bias. We believe, however, that it is unlikely that our observations were generated or distorted by a systematic bias in the parental responses to the questions about their child’s health or swimming practice. The strongest argument against that possibility is that the most statistically significant observations were made with objective outcome measures (serum pneumoproteins and EIB test) that parents were not aware of when filling the questionnaires. The parents were also blinded to the tested hypothesis because initially the study was not designed to look specifically at the effects of infant swimming practice. In addition, the tested hypothesis of adverse respiratory effects associated with infant swimming is far from being common in the community. Actually, this hypothesis is just the opposite of the belief that the parents probably have when they decide to take their infant to swimming pool. The possibility that they
could have biased their responses in favor of adverse effects seems thus unlikely.

Our data suggest that the poorer respiratory health of children who swam as infants could be linked to distal airways damage detected by the assay of serum pneumo-proteins. The concentration of serum CC16 is a well-validated marker of the lung epithelium barrier integrity, reflecting either the number of Clara cells lining terminal airways or the permeability of the alveolar-capillary barrier.26,27 When adjusted for the level of a surfactant-associated protein, such as SP-D, the concentration of CC16 proves to be an even more sensitive marker, probably because this ratio integrates both the damage to Clara cells and the increased leakiness of the alveolar-blood barrier.30 A decrease in serum CC16 reflecting a parallel loss of Clara cells was demonstrated in humans and in rodents acutely or chronically exposed to a variety of lung toxicants.31 For instance, active smoking leads to a 20% to 30% decrease of serum CC16, mirroring a parallel decrease in the number of Clara cells.30–34 A similar decrease of serum CC16 was observed after occupational exposures to crystalline silica or firesmoke.35 The decrease of serum CC16 in children who swam as infants averaged 20%. Thus, it is almost of the same magnitude as that caused in adults by tobacco smoke or industrial chemicals. If one refers to the clinical consequences of active smoking or of occupational exposures to crystalline silica and other lung irritants, it would not be surprising that similar alterations of the respiratory epithelium in young children could make them more prone to develop some respiratory diseases. This interpretation is also consistent with the antiinflammatory properties of CC16,37,38 as well as with the finding that asthma and other respiratory diseases40 are associated with lower intrapulmonary pools of CC16 because of Clara cell damage or the intravascular leakage of the protein.

Given the lack of data concerning the toxicity of

---

**FIGURE 1**

Effects of total asthma (A), recurrent bronchitis (B), and swimming infant practice (A and B) and of their possible interactions on the serum concentrations of CC16 and SP-D and the serum CC16/SP-D ratio. Results were assessed by 2-way analysis of variance followed by Dunett’s multiple-comparison test. Mean (with SE) values that are significantly different from that of controls: • $P < .05$, • $P < .01$. 

---

1100 BERNARD et al

**BERNARD et al**
swimming pool chemicals for the respiratory tract of infants and young children, the pool factor responsible for these airways alterations is difficult to identify. We strongly suspect, however, that the chlorination products that infants inhale as gases, aerosols, or even water repeatedly damage the airway epithelium. One culprit might be trichloramine, the highly volatile and reactive gas formed when chlorine reacts with organic matter brought by swimmers. This gas that gives indoor swimming pools their characteristic chlorine smell was, indeed, found to cause asthma and pulmonary epithelium damage in lifeguards and recreational swimmers. Because their lungs are still developing, infants could be particularly sensitive to this irritant and be affected despite the rather limited time they spend in pools (usually 20–30 minutes per session). The levels of trichloramine in the studied pool were on average below the provisional 2-hour air quality guideline of 500 μg/m³ recently recommended by the World Health Organization. If trichloramine is responsible for the respiratory effects observed in our study, this would mean that the World Health Organization guideline is too high and should be lowered to be more in accordance with studies showing that trichloramine can disrupt the lung epithelium barrier of swimmers at concentrations ranging from 355 to 490 μg/m³. Another important route of exposure that might cause significant damage to infant’s lungs is the inhalation of aerosols or of small volumes of chlorinated water when the infants actively play or have their head under water. Damage to the respiratory tract caused by the inhalation of heavily chlorinated water seems especially plausible because the maximum recommended levels for chlorine in the Brussels pool that our children attended when they were infants were relatively high. For instance, the recommended limit for combined chlorine was 2 ppm, a concentration twice higher than the current standard (0.8 ppm) and 10 times higher than the standard applied in Germany (0.2 ppm). Because infants cannot control their breathing as well as when they get older, the infant swimming practice is considered to be safe because of the laryngeal or gag reflex that is triggered when water gets into infant’s mouth. By closing off the larynx with the epiglottis, this reflex is supposed to keep pool water from entering the lungs. However, the gag reflex, even if very effective, cannot prevent small amounts of chlorinated water deposited or trapped in the upper respiratory tract to be conducted more deeply in the lungs when the infant surfaces to breathe. Cases of hyponatremic described after dunking infants in swimming pools attest to the amount of water that an infant can ingest and possibly inhale when being under water. Last, one cannot exclude the possibility that the inhalation of hypotonic water also causes some epithelial changes aggravating the effects of chlorination products.

Although we did not specifically interview the parents about this issue, our study provides some interesting insights into the reasons why parents take their infant to swimming programs. As one could expect, prevention of drowning seems to be an important reason because there were 5 times more children having access to a backyard pool among the infant swimming group than among the other children. Although drowning is a leading cause of unintentional injury and death in children, frequently involving backyard pools, it should be noted that the protection offered by infant swimming programs is much debated. According to the American Academy of Pediatrics, there is indeed no clearcut scientific evidence that the likelihood of drowning is reduced by the participation to such programs. The American Academy of Pediatrics argues that until the age of 4 years, infants are not developmentally ready for formal swimming lessons and that the participation in such programs could give to parents a false sense of security about their child’s skill in water. The American Academy of Pediatrics did not consider chemical hazards, but if, as suggested by our findings, chlorine used to disinfect pools poses some threat to infants’ health, this would certainly further justify a critical appraisal of infant swimming programs. There might be, however, an even more debatable reason encouraging parents to participate to swimming infant programs. Intriguingly, the infant swimming group included a much greater proportion of children who had been exposed to tobacco smoke in utero or at home. The origin of this difference is unclear, but we think it might reflect the will of the parents to mitigate for the adverse effects of passive smoking by giving their infant the possibility to practice a healthy activity. The unfortunate irony is that the result might well be just the opposite, because our study suggests that infant swimming activity interacts with passive smoking to greatly increase the risk of asthma or recurrent bronchitis.

The major limitation of our study certainly lies in the relatively small size of the infant swimming group. With <15% of Brussels schoolchildren participating in such programs, this was the maximum sample size we could achieve in a population-based study involving 341 children. The alternative would be to recruit children via infant swimming organizations. However, assuming that these organizations would be willing to collaborate in research exploring the risks linked to chlorinated pools, the study would then be confronted with the difficulty of avoiding response or selection bias and, above all, with the difficulty of recruiting a well-matched control population. Fortunately, the limited statistical power of our study was balanced by the use of sensitive outcome measures, leading to particular strong associations that were unlikely to be explained by chance only. The small size of our study, however, did not enable us to draw definitive conclusions regarding the impact of backyard pool chlorine to which children in the infant swimming
CONCLUSIONS
Our study shows that the infant swimming practice in indoor chlorinated pools can be associated with airways changes that predispose children to asthma and recurrent bronchitis later in childhood. Given the increasing popularity of swimming pools, there is a definitive need to assess the effects of chlorination products on the respiratory tract of very young children. In the meantime, because in most countries these products are not as strictly regulated and monitored as indicators of microbiological risks, we can only recommend caution before regularly taking infants to poorly maintained swimming pools with excessive levels of chlorine in the water and in the air.

ACKNOWLEDGMENTS
This study was supported by the Brussels Capital Region. Dr Bernard is research director of the National Fund for Scientific Research in Belgium.

REFERENCES


Mattsson J, Remberger M, Andersson O, Sundberg B, Nord M. Decreased serum levels of Clara cell secretory protein (CC16) are associated with bronchiolitis obliterans and may permit early diagnosis in patients after allogeneic stem-cell transplantation. Transplantation. 2005;79:1411–1416


Freedman FB. Water Babies. London, United Kingdom: Lorenz Books; 2003


THE MOMMY TRACK

“Why do women fall off academia’s science track at a faster clip than men? The cause is not innate sex differences, a new study suggests, but neither is it a simple matter of gender discrimination. If a problem exists, the authors conclude, it’s about motherhood, not women in general. Using the 1973–2001 Survey of Doctorate Recipients, the study found that while women are less likely than men to enter tenure-track positions in the sciences, the difference is explained completely by ‘fertility decisions.’ Single, childless women are between 11 percent (in the life sciences) and 21 percent (in the physical sciences) more likely to have a tenure-track job within five years of finishing their doctorate than single, childless men. . . . Children can hurt a woman’s chances significantly: Having a child at pre-kindergarten age took 8 percentage points off a woman’s chance of getting that tenure-track job.”

Ginther DK. Atlantic. March 2007
Noted by JFL, MD
Maternal Asthma and Maternal Smoking Are Associated With Increased Risk of Bronchiolitis During Infancy

Kecia N. Carroll, MD, MPH,a,b,c Tebeb Gebretsadik, MPHd, Marie R. Griffin, MD, MPH,e,f,g,h,i, William D. Dupont, PhDd,f, Edward F. Mitchel, MSf, Pingsheng Wu, PhDh,j, Rachel Enriquez, RN, PhDj, Tina V. Hartert, MD, MPH,e,j,k,l,m

Departments of aPediatrics, cMedicine, fPreventive Medicine, and gBiostatistics, Divisions of aGeneral Pediatrics, fGeneral Internal Medicine, and iAllergy, Pulmonary and Critical Care Medicine, cChild and Adolescent Health Research Unit, fCenter for Education and Research on Therapeutics, fCenter for Health Services Research, and iGeneral Clinical Research Center, Vanderbilt University School of Medicine, Nashville, Tennessee; dMid-South Geriatric Research Education and Clinical Center and Clinical Research Center of Excellence, Veterans Affairs Tennessee Valley Health Care System, Nashville, Tennessee; jMeharry/Vanderbilt Center for Reducing Asthma Disparities, Nashville, Tennessee

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

ABSTRACT

OBJECTIVE. Our goal was to determine whether maternal asthma and maternal smoking during pregnancy are associated with the incidence and severity of clinically significant bronchiolitis in term, otherwise healthy infants without the confounding factors of small lung size or underlying cardiac or pulmonary disease.

PATIENTS AND METHODS. We conducted a population-based retrospective cohort study of term, non–low birth weight infants enrolled in the Tennessee Medicaid Program from 1995 to 2003. The cohort of infants was followed through the first year of life to determine the incidence and severity of bronchiolitis as determined by health care visits and prolonged hospitalization.

RESULTS. A total of 101 245 infants were included. Overall, 20% of infants had ≥1 health care visit for bronchiolitis. Compared with infants with neither factor, the risk of bronchiolitis was increased in infants with maternal smoking only, maternal asthma only, or both. Infants with maternal asthma only or with both maternal smoking and asthma had the highest risks for emergency department visits and hospitalizations. Infants with a mother with asthma had the highest risk of a hospitalization >3 days, followed by infants with both maternal asthma and smoking, and maternal smoking only.

CONCLUSIONS. Maternal asthma and maternal smoking during pregnancy are independently associated with the development of bronchiolitis in term, non–low birth weight infants without preexisting cardiac or pulmonary disease. The risk of bronchiolitis among infants with mothers who both have asthma and smoke during pregnancy is ~50% greater than that of infants with neither risk factor. Efforts to decrease the illness associated with these 2 risk factors will lead to decreased morbidity from bronchiolitis, the leading cause of hospitalization for severe lower respiratory tract infections during infancy.
Bronchiolitis, caused by viruses such as respiratory syncytial virus (RSV) and rhinovirus, is the leading cause of lower respiratory tract infections in infants. Bronchiolitis results in significant morbidity in infants, accounting for an estimated 120,000 hospitalizations annually in the United States among children <1 year of age. It is well established that children with medical conditions such as prematurity, chronic lung disease, and congenital heart disease are at an increased risk of hospitalization for bronchiolitis. Risk factors implicated in the development of severe bronchiolitis include young age, birth early in the RSV season, male gender, lack of breastfeeding, and passive smoke exposure. However, 50% of hospitalizations for bronchiolitis involve term or near-term, otherwise healthy infants, among whom there are few established risk factors for bronchiolitis.

Viruses such as RSV infect the majority of children during infancy; however, only 20% to 30% develop bronchiolitis and only ~3% are hospitalized. Infants who develop severe bronchiolitis have an increased risk of recurrent wheezing and/or asthma later in life. It is unclear whether a predisposition to develop asthma is a risk factor for the development of bronchiolitis or whether the viral infection is involved in asthma inception. Numerous studies, which typically had small sample sizes, have looked for an association between asthma in a first-degree relative and severe bronchiolitis in early life. However, the majority of studies did not find an association. Maternal smoking during pregnancy and/or other environmental tobacco smoke exposure (ETS) have been associated with bronchiolitis severity. However, there have been no large population-based studies examining the interaction of factors, such as maternal smoking during pregnancy and maternal asthma, on the development of bronchiolitis in term, healthy infants without the confounding factors of preexisting heart or lung disease. The objectives of this study were to estimate the association of maternal asthma and maternal smoking during pregnancy with clinically significant bronchiolitis in a large population of term, otherwise healthy infants enrolled in the Tennessee Medicaid program during 1995–2003. We hypothesized that maternal smoking during pregnancy and maternal asthma are associated with the incidence and severity of bronchiolitis during infancy independently of preexisting pulmonary or cardiac disease.

**PATIENTS AND METHODS**

We conducted a population-based retrospective cohort study of >100,000 women and infants enrolled in the Tennessee Medicaid Program, during 1995–2003. In 1994, the Tennessee Medicaid Program replaced the federal Medicaid program as a state-based managed health care program that covered Medicaid-eligible individuals and the uninsured. Approximately 50% of infants born in Tennessee are born to women enrolled in the Tennessee Medicaid Program. Infants born to women of black or white race, aged 15 to 44 years, and continuously enrolled in the Tennessee Medicaid program were eligible for study inclusion. Infants of mothers of other races were too few to study, therefore, 2.13% of otherwise eligible infants were excluded. Infants of women with unknown smoking history (0.2%) were also excluded. We defined continuous enrollment from the year before pregnancy (last menstrual period date through 365 days) through the date of delivery with no more than 45 days of nonenrollment. To overcome potential confounding associated with preexisting cardiac or pulmonary disease, we included only healthy infants born at ≥37 weeks’ estimated gestational age (EGA) with birth weight of ≥2500 g. The protocol was approved by the institutional review boards of Vanderbilt University and the Tennessee Department of Health.

Using previously described methods, we obtained study data from linked Tennessee Medicaid administrative data files and Tennessee State vital records. Analysis files contained no personal identifiers, and study results are reported in aggregate and cannot be linked to individuals. We determined EGA in weeks by the date of the last menstrual period on the birth certificate (91.7%), or calculated based on the median gestational period in weeks for the infant’s race, birth weight, and birth year (8.22%) or assigned last menstrual period as 270 days before birth (0.04%). Infants with any of the following during the first 3 months of life were excluded: an International Classification of Diseases, Ninth Revision (ICD-9) code or Current Procedural Terminology code indicating congenital heart disease or surgery for congenital heart disease respectively, an ICD-9 code indicating chronic lung disease or a congenital anomaly of the upper airway and/or esophagus, or receipt of ≥1 dose of RSV prophylaxis (Palivizumab or RSV immune globulin).

Using a previously validated method during the predelivery eligibility period we identified maternal asthma through health encounter and pharmacy claims. Women with an ICD-9 code of 493 (asthma) in any of the 9 diagnostic fields for inpatient, other hospital care (23 hour observation), or outpatient physician visit claims were considered to have asthma. In addition, women with 2 prescriptions for any short-acting β-agonist or a single prescription for any other asthma medication (long-acting β agonist, inhaled corticosteroids, leukotriene modifying agents) were considered to have asthma. Demographic characteristics identified from Tennessee Medicaid enrollment files included self-reported maternal race (black, white) and region of residence (urban, suburban, rural). Women of nonblack or nonwhite race were too few to study. Demographic variables determined from infant birth certificate data included self-reported maternal smoking during preg-
nancy, maternal age at delivery, maternal education level, marital status (single, married), siblings (none, 1, ≥2), and infant gender. We determined infant health care visits for bronchiolitis year-round using ICD-9 codes for bronchiolitis (466.1) and/or RSV pneumonia (480.1).

To determine severity of bronchiolitis, we applied a hierarchy: infants with any hospitalization were categorized in the hospitalization group, infants with an emergency department (ED) visit and no hospitalization were categorized in the ED group, infants with only clinic visits were classified in the clinic group, and infants without any health care visits for bronchiolitis were categorized in the no visit group. We also determined the risk of a prolonged hospital stay defined by >3-day length of stay, based on previous reports of median length of stay for bronchiolitis for children in the United States of 3 days.2

The main predictor variables were history of maternal asthma and maternal smoking during pregnancy. Descriptive statistics for categorical variables were expressed as proportions and for continuous variables as median and interquartile range for nonnormally distributed variables. All infants in the study cohort were followed until 1 year of age, until they had >21 days of nonenrollment in Tennessee Medicaid, or death. To account for age at time of diagnosis, we used survival analyses with age as the time-dependent variable. We used Kaplan-Meier curves to estimate the cumulative incidence of a bronchiolitis health care visit (any clinic, ED, 23-hour observation, and/or hospitalization). We compared the cumulative incidence of clinically significant bronchiolitis by maternal asthma and smoking status using the 2-sided log-rank test. We obtained the relative risk of a bronchiolitis diagnosis using Cox’s proportional hazards regression.36 We included variables in the model based on their clinical importance and association with maternal asthma, maternal smoking, and bronchiolitis including: maternal smoking during pregnancy and maternal asthma status, region of residence, maternal age, maternal education level, maternal race/ethnicity, other living siblings, infant birth weight, and gender.5,37 We assessed for interaction between maternal asthma and maternal smoking by including a cross-product in the model. Using a multinomial logistic regression model, we estimated the risk of having a clinic visit, ED visit, or hospitalization (including 23-hour observations) compared with none by maternal asthma and smoking status. We estimated the population-based cumulative incidence of a prolonged hospitalization (>3 days) for bronchiolitis using survival analyses and obtained the relative risk of a prolonged hospitalization using Cox’s proportional hazards regression. We confirmed that the proportional hazards assumption was reasonable for our data using log-log plots. R-software 2.11 (www.r-project.org), SAS 8.2 (SAS Institute, Cary, NC) and Stata 8.2 were used for data analyses. We used a 2-sided 5% significance level for all statistical inferences.

RESULTS

A total of 101 245 mother-infant dyads were included in the study: 42% of the mothers were black. The median age of women at delivery was 22 years, and 65% of women were single. Forty-two percent of women had less than a high school degree. The women were from urban (44%), suburban (23%), and rural regions (33%) of the state. Approximately 7% of women met the criteria for asthma during the predelivery baseline period, and 28% of women reported smoking during pregnancy. The majority of women were nonsmokers and did not have asthma (67.9%), whereas 25.3% of women smoked during pregnancy but did not have asthma, 4.4% of women had asthma and did not smoke, and 2.5% of women both smoked during pregnancy and had a diagnosis of asthma (Table 1). Among the infants, 51% were male, the median EGA was 39.6 weeks, and the median birth weight was 3289 g. Approximately 29% of infants had no living siblings, 37% had 1 sibling, and 34% had ≥2 siblings.

Overall, 20% of infants had at least 1 bronchiolitis visit (clinic visit, ED visit, and/or hospitalization for bronchiolitis). Differences were seen by maternal asthma and smoking status in the frequency of having a bronchiolitis diagnosis during infancy. The proportions of infants with a bronchiolitis diagnosis in the first year of life were 18% in children with neither maternal asthma nor smoking, 24% among those with maternal asthma only, 24% among those with maternal smoking only, or 30% among those with both smoking and asthma, respectively. Differences in the unadjusted incidence of bronchiolitis are illustrated by Kaplan-Meier curves (Fig 1). Infants with both risk factors had the greatest risk of a health care encounter for bronchiolitis (hazard ratio [HR]: 1.47; 95% confidence interval [CI]: 1.36–1.59), adjusted for region of residence in the state, maternal age, maternal education level, maternal race, other living siblings, infant birth weight, and gender, followed by infants with maternal asthma only (adjusted HR: 1.39; 95% CI: 1.30–1.48), and maternal smoking only (adjusted HR: 1.14; 95% CI: 1.10–1.18), compared with infants with neither risk factor (Table 2).

The risks of clinic visits, ED visits, or hospitalizations for bronchiolitis among infants whose mothers did or did not smoke during their pregnancy or have asthma are given in Table 3. Infants whose mothers both smoked and had asthma were more likely to have ED visits for bronchiolitis than infants with neither of these risk factors (adjusted odds ratio [OR]: 2.18; 95% CI: 1.87–2.54). Infants whose mothers had only 1 of these risk factors had a lower but still significantly elevated risk for ED visits. Infants whose mothers either had asthma or smoked were at significantly increased risk for clinic
TABLE 1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
</tr>
<tr>
<td>Maternal age, median (IQR), yb</td>
<td>22 (19–25)</td>
<td>22 (20–26)</td>
<td>22 (19–25)</td>
<td>22 (20–26)</td>
</tr>
<tr>
<td>Maternal race</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>33 108 (48.19)</td>
<td>21 451 (83.79)</td>
<td>2431 (54.91)</td>
<td>2215 (88.35)</td>
</tr>
<tr>
<td>Black</td>
<td>35 602 (51.81)</td>
<td>4150 (16.21)</td>
<td>1996 (45.09)</td>
<td>292 (11.65)</td>
</tr>
<tr>
<td>Education, y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;12</td>
<td>26 184 (38.18)</td>
<td>13 319 (52.13)</td>
<td>1887 (42.71)</td>
<td>1318 (52.72)</td>
</tr>
<tr>
<td>12</td>
<td>31 991 (46.65)</td>
<td>10 276 (40.22)</td>
<td>1925 (43.57)</td>
<td>986 (39.44)</td>
</tr>
<tr>
<td>&gt;12</td>
<td>10 406 (15.17)</td>
<td>1 955 (7.65)</td>
<td>606 (13.72)</td>
<td>196 (7.84)</td>
</tr>
<tr>
<td>Unknown</td>
<td>129 (0.19)</td>
<td>51 (0.20)</td>
<td>9 (0.20)</td>
<td>7 (0.28)</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>45 623 (66.40)</td>
<td>15 429 (60.28)</td>
<td>2826 (63.84)</td>
<td>1457 (58.12)</td>
</tr>
<tr>
<td>Married</td>
<td>23 084 (33.60)</td>
<td>10 166 (39.72)</td>
<td>1601 (36.16)</td>
<td>1050 (41.88)</td>
</tr>
<tr>
<td>Unknown</td>
<td>3 (0.00)</td>
<td>6 (0.00)</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>Region of residence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>35 416 (51.54)</td>
<td>6838 (26.71)</td>
<td>2112 (47.71)</td>
<td>608 (24.25)</td>
</tr>
<tr>
<td>Suburban</td>
<td>13 341 (19.71)</td>
<td>7508 (29.33)</td>
<td>978 (22.09)</td>
<td>721 (28.76)</td>
</tr>
<tr>
<td>Rural</td>
<td>19 689 (28.66)</td>
<td>11 219 (43.82)</td>
<td>1335 (30.16)</td>
<td>1178 (46.99)</td>
</tr>
<tr>
<td>Unknown</td>
<td>64 (0.09)</td>
<td>36 (0.14)</td>
<td>2 (0.05)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>Siblings</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>20 671 (30.11)</td>
<td>6680 (26.11)</td>
<td>1448 (32.75)</td>
<td>736 (29.38)</td>
</tr>
<tr>
<td>1</td>
<td>24 876 (36.23)</td>
<td>9717 (37.98)</td>
<td>1495 (33.81)</td>
<td>900 (35.93)</td>
</tr>
<tr>
<td>&gt;2</td>
<td>23 115 (33.66)</td>
<td>9190 (35.92)</td>
<td>1479 (33.45)</td>
<td>869 (34.69)</td>
</tr>
<tr>
<td>Unknown</td>
<td>48 (0.07)</td>
<td>14 (0.05)</td>
<td>5 (0.11)</td>
<td>2 (0.08)</td>
</tr>
<tr>
<td>Infant gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>35 150 (51.16)</td>
<td>13 215 (51.62)</td>
<td>2288 (51.68)</td>
<td>1352 (53.93)</td>
</tr>
<tr>
<td>Female</td>
<td>33 560 (48.84)</td>
<td>12 386 (48.38)</td>
<td>2139 (48.32)</td>
<td>1155 (46.07)</td>
</tr>
<tr>
<td>Birth weight, g</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (IQR)</td>
<td>3317 (3033–3629)</td>
<td>3203 (2920–3487)</td>
<td>3317 (3033–3629)</td>
<td>3175 (2920–3487)</td>
</tr>
<tr>
<td>Infant EGA, wk</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median (IQR)</td>
<td>39.6 (38.6–40.7)</td>
<td>39.6 (38.6–40.7)</td>
<td>39.6 (38.6–40.7)</td>
<td>39.6 (38.6–40.7)</td>
</tr>
</tbody>
</table>

IQR indicates interquartile range.

* Both maternal smoking during pregnancy and maternal asthma.

# Maternal age at delivery.

FIGURE 1
visits and/or hospitalizations. Maternal asthma was the more important of these 2 risk factors and was associated with ORs of 1.20 and 1.74 for clinic visits and hospitalizations, respectively. The combined effects of both of these risk factors were similar to those for maternal asthma alone.

Table 4 gives the risks of a prolonged hospitalization for bronchiolitis classified by maternal asthma and smoking status. The median length of stay in our study was 3 days. Overall, 2.26% of the over 100,000 infants in the cohort had a hospital stay of >3 days. Figure 2 illustrates that infants with 1 or both risk factors were at greater risk of having a hospital stay >3 days for bronchiolitis than were infants with neither risk factor. Compared with infants without maternal asthma or smoking, infants with maternal asthma had the highest risk of having a hospitalization >3 days (HR: 1.52; 95% CI: 1.26–1.82), followed by infants with both maternal asthma and smoking (HR: 1.38; 95% CI: 1.12–1.71), and maternal smoking only (HR: 1.19; 95% CI: 1.08–1.31; Table 4). Differences in the trend in the HRs from the trend in the cumulative incidences were because of covariate adjustment (data not shown).

**DISCUSSION**

In this population-based cohort of >100,000 term, healthy infants, we determined that maternal asthma, which represents a predisposition to have children who develop asthma,38 and maternal smoking during pregnancy are independent risk factors for developing clinically significant bronchiolitis in the first year of life. Overall, 20% of infants had an outpatient or inpatient visit for bronchiolitis, which is consistent with previous reports of bronchiolitis incidence.10 The risk of bronchiolitis was significantly increased in patients with maternal smoking or maternal asthma, with the highest risk seen in patients with both maternal asthma and materi-
nal smoking during pregnancy. The increased risk seen in infants with both risk factors highlights the clinical importance of the combined effects of these risk factors on the development of bronchiolitis during infancy.

Importantly, this study included only term, non–low birth weight children without preexisting heart or lung disease. We also captured both outpatient and inpatient care for bronchiolitis. Therefore, we were able to investigate the impact of a familial predisposition to develop asthma and maternal smoking during pregnancy on the spectrum of bronchiolitis severity in a population of infants without the confounding of small lung size, or underlying pulmonary or cardiac disease.

Maternal asthma, maternal smoking during pregnancy, or both also increased the risk of severe bronchiolitis in this population of term, healthy infants. Infants with 1 or both risk factors were more likely to have an ED visit or hospitalization compared with infants with neither factor. Hospitalization, in particular, is a standard measure of severe bronchiolitis used by scientists for >40 years to conduct epidemiologic studies of bronchiolitis. Although most infants will become infected with RSV and 20% to 30% will develop bronchiolitis, only a small percentage of infected infants will be hospitalized. In our study cohort, infants with a mother with asthma or both maternal asthma and smoking had the highest risks of being hospitalized for bronchiolitis and having a prolonged hospitalization for bronchiolitis.

Having a mother with asthma is an easily identifiable risk factor for the development of bronchiolitis. Clarifying the association between a familial predisposition to develop asthma and bronchiolitis will guide immunologic and mechanistic studies to gain insight into differential host responses to viral infection during infancy and help define high-risk target populations for preventive efforts. In addition, providers can inform families of increased risk and further encourage the importance of avoidance measures to reduce the risk of viral infections. Regarding future interventions, several RSV vaccines are being evaluated, and if approved, could be used to prevent disease, particularly in high-risk children.

Maternal smoking during pregnancy is a preventable factor in bronchiolitis morbidity to which >25% of our study cohort was exposed. We focused only on non–low birth weight infants in this study and found that infants with maternal smoking during pregnancy have an increased risk of developing clinically significant bronchiolitis and of more severe bronchiolitis independent of the effects of low birth weight. The increased incidence and severity of bronchiolitis in infants with maternal smoking highlights the importance of efforts to prevent tobacco use and promote smoking cessation in pregnant women and women of childbearing age.

There are several potential limitations of this work. We defined maternal asthma using ICD-9 diagnosis codes and medication use and thus may not detect individuals with intermittent disease. However, the asthma prevalence in the study cohort was similar to the reported asthma prevalence in Tennessee. Our definition of asthma has been demonstrated to be specific, with previous work showing that similarly defined individuals had definite asthma (62%) or probable asthma (38%) by chart review. We determined 1 measure of a genetic predisposition to develop asthma, namely maternal asthma; however, we were not able to obtain the asthma history of other first-degree relatives. Therefore, there are likely individuals in the group of infants...
without maternal asthma who have a familial predisposition to develop asthma. This misclassification would lead to an attenuation of the association between familial predisposition to develop asthma and bronchiolitis. This suggests, for example that the 74% increased odds for infants with maternal asthma to be hospitalized for bronchiolitis compared with infants without maternal asthma or smoking is a real difference and potentially an underestimation of the impact of familial predisposition. One might also question whether assessing maternal smoking from birth certificates would be accurate. However, previous reports from our research group have demonstrated good concordance of the birth certificate data with medical chart data. In addition, maternal smoking history is determined before the infant’s birth and any misclassification would be nondifferential for those infants that develop subsequent respiratory symptoms. We did not have access to the smoking history of others in the infants’ households nor could we separate in utero tobacco smoke exposure from maternal smoking during the postnatal period. Therefore, maternal smoking during pregnancy is likely a measure of prenatal and postnatal ETS. In this case, the association of ETS and bronchiolitis during infancy would be underestimated because there will be misclassification. We identified bronchiolitis by using administrative data, and ICD-9 diagnoses represent objective physician-characterized outcomes at the time of illness that would not be influenced by recall bias. We used ED visits as a measure of bronchiolitis severity, and it is possible that some infants were taken to the ED that could have been cared for in outpatient clinics. However, the increased risk of an ED visit was consistent with other measures of bronchiolitis severity including hospitalization and prolonged length of stay. Because of the retrospective nature of this cohort study, it is possible that study findings were influenced by other unmeasured factors. In addition, we cannot rule completely that we have eliminated all possible bias in our multivariable statistical analyses. Although the generalizability of our findings to non-Medicaid populations may be questioned, half of the infants born in Tennessee are enrolled in the Medicaid program. Therefore, infants in Tennessee Medicaid are highly representative of the state’s population and a substantial proportion of children born in other areas in the United States.

CONCLUSIONS

In a population-based cohort of >100,000 healthy, non-low birth weight infants, maternal asthma and maternal smoking during pregnancy were both associated with an increased incidence and severity of bronchiolitis during infancy. Study findings suggest that a familial predisposition to develop asthma and maternal smoking during pregnancy are independent risk factors for the development of the leading cause of lower respiratory tract infections during infancy. Because viruses such as RSV are ubiquitous, the association of severe bronchiolitis and maternal asthma suggests that genetic/host factors influence infant response to viral infection and confer increased risk for viral lower respiratory tract infections. The work in understanding differences in the immune response to viral infections will increase our understanding of the etiology of the increased illness experienced by children with maternal asthma. In addition, the study findings support future investigations aimed at decreasing bronchiolitis severity in high-risk populations by targeting infants whose mothers smoke during pregnancy and/or have a familial predisposition to develop asthma.

ACKNOWLEDGMENTS

This study was supported by the National Institutes of Health (grants U01 HL 72471, MO1 RR00095, KO8 A101582, and K12 RR17697); the Agency for Healthcare Research and Quality, Centers for Education and Research (grant U18-HS10384); the Geriatric Research Education and Clinical Center, Department of Veterans Affairs; and the Thrasher Research Fund.

We thank the Tennessee Bureau of TennCare (Department of Finance and Administration) and the Tennessee Department of Health (Office of Policy, Planning and Assessment) for providing the data.

REFERENCES

18. Gern JE. Viral respiratory infection and the link to asthma. *Pediatr Infect Dis J.* 2004;23(suppl 1):S78–S86
51. Piper JM, Mitchel EF Jr, Snowden M, Hall C, Adams M, Taylor...
TYMPANOSTOMY TUBES AND DEVELOPMENTAL OUTCOMES AT 9 TO 11 YEARS OF AGE

**Background:** Developmental impairments in children have been attributed to persistent middle-ear effusion in their early years of life. Previously, we reported that among children younger than 3 years of age with persistent middle-ear effusion, prompt as compared with delayed insertion of tympanostomy tubes did not result in improved cognitive, language, speech, or psychosocial development at 3, 4, or 6 years of age. However, other important components of development could not be assessed until the children were older.

**Methods:** We enrolled 6350 infants soon after birth and evaluated them regularly for middle-ear effusion. Before 3 years of age, 429 children with persistent effusion were randomly assigned to undergo the insertion of tympanostomy tubes either promptly or up to 9 months later if effusion persisted. We assessed literacy, attention, social skills, and academic achievement in 391 of these children at 9 to 11 years of age.

**Results:** Mean (±SD) scores on 48 developmental measures in the group of children who were assigned to undergo early insertion of tympanostomy tubes did not differ significantly from the scores in the group that was assigned to undergo delayed insertion. These measures included the Passage Comprehension subtest of the Woodcock Reading Mastery Tests (mean score, 98±12 in the early-treatment group and 99±12 in the delayed-treatment group); the Spelling, Writing Samples, and Calculation subtests of the Woodcock-Johnson III Tests of Achievement (96±13 and 97±16; 104±14 and 105±15; and 99±13 and 99±13, respectively); and inattention ratings on visual and auditory continuous performance tests.

**Conclusions:** In otherwise healthy young children who have persistent middle-ear effusion, as defined in our study, prompt insertion of tympanostomy tubes does not improve developmental outcomes up to 9 to 11 years of age. (ClinicalTrials.gov number, NCT00365092.)
Laboratory Values for Children With Newly Diagnosed Inflammatory Bowel Disease

David R. Mack, MD, Christine Langton, MSW, MPH, James Markowitz, MD, Neal LeLeiko, MD, PhD, Anne Griffiths, MD, Athos Bousvaros, MD, Jonathan Evans, MD, Subra Kugathasan, MD, Anthony Otley, MD, Mariann Pfefferkorn, MD, Joel Rosh, MD, Adam Mezoff, MD, Susan Moyer, MD, Maria Oliva-Hemker, MD, Robert Rothbaum, MD, Robert Wylie, MD, J. Fernando delRosario, MD, David Keljo, MD, PhD, Trudy Lerer, MS, Jeffrey Hyams, MD, for the Pediatric Inflammatory Bowel Disease Collaborative Research Group

OBJECTIVE. The goal was to determine how often common laboratory tests yield normal results at the time of diagnosis for children with inflammatory bowel disease.

METHODS. Data were obtained from a registry of children with newly diagnosed inflammatory bowel disease who were enrolled prospectively in 18 US/Canadian centers. Laboratory values investigated included hemoglobin level, platelet count, albumin level, and erythrocyte sedimentation rate. Disease severity was categorized by physician global assessment.

RESULTS. A total of 526 children (mean age: 11.6 years; 58% male; 392 with Crohn disease and 134 with ulcerative colitis) were studied. All 4 values were normal for 21% of patients with mild Crohn disease and 54% with mild ulcerative colitis. In contrast, only 3.8% of children with moderate/severe Crohn disease and 4.3% with moderate/severe ulcerative colitis had normal results for all 4 tests. The erythrocyte sedimentation rate was least likely to be normal; overall, 26% of patients with inflammatory bowel disease had a normal erythrocyte sedimentation rate, including 18% with moderate/severe disease. Hemoglobin levels were normal for 32%, platelet counts for 50%, and albumin levels for 60%. There was no clear association between Crohn disease location and either severity or number of normal laboratory values. In contrast, there were direct correlations between ulcerative colitis disease severity and both the extent of bowel inflammation and the number of abnormal laboratory tests.

CONCLUSION. The presence of normal screening laboratory studies should not dissuade clinicians from considering a diagnosis of inflammatory bowel disease.

ABSTRACT

Objective. The goal was to determine how often common laboratory tests yield normal results at the time of diagnosis for children with inflammatory bowel disease.

Methods. Data were obtained from a registry of children with newly diagnosed inflammatory bowel disease who were enrolled prospectively in 18 US/Canadian centers. Laboratory values investigated included hemoglobin level, platelet count, albumin level, and erythrocyte sedimentation rate. Disease severity was categorized by physician global assessment.

Results. A total of 526 children (mean age: 11.6 years; 58% male; 392 with Crohn disease and 134 with ulcerative colitis) were studied. All 4 values were normal for 21% of patients with mild Crohn disease and 54% with mild ulcerative colitis. In contrast, only 3.8% of children with moderate/severe Crohn disease and 4.3% with moderate/severe ulcerative colitis had normal results for all 4 tests. The erythrocyte sedimentation rate was least likely to be normal; overall, 26% of patients with inflammatory bowel disease had a normal erythrocyte sedimentation rate, including 18% with moderate/severe disease. Hemoglobin levels were normal for 32%, platelet counts for 50%, and albumin levels for 60%. There was no clear association between Crohn disease location and either severity or number of normal laboratory values. In contrast, there were direct correlations between ulcerative colitis disease severity and both the extent of bowel inflammation and the number of abnormal laboratory tests.

Conclusion. The presence of normal screening laboratory studies should not dissuade clinicians from considering a diagnosis of inflammatory bowel disease.
LABORATORY EVALUATION TO screen for evidence of inflammatory bowel disease (IBD) is performed routinely for children with chronic abdominal pain or diarrhea and can help both to establish a diagnosis and to serve as a baseline for management. When these children are ill with concomitant bleeding, weight loss, and abdominal tenderness, generally prompt referral to a pediatric gastroenterologist is made. For children with milder symptoms, such as only occasional abdominal discomfort or intermittent episodes of loose stools, the presence of normal laboratory parameters might reassure the primary health care provider that IBD is not present. By using a large, prospective, pediatric IBD database, we sought to identify the frequency with which children with newly diagnosed IBD present with normal laboratory values in the tests commonly used to screen for IBD.

METHODS

Patients

All data included in this report were obtained from the database of the Pediatric IBD Collaborative Research Group Registry. This registry was initiated in January 2002 by 18 US and Canadian pediatric gastroenterology centers with significant clinical expertise in IBD, as a method of describing the contemporary natural history of IBD in patients with newly diagnosed disease who had not reached their 16th birthday. Diagnoses of Crohn disease (CD) and ulcerative colitis (UC) were made at each of the 18 participating centers on the basis of conventional clinical, laboratory, radiologic, endoscopic, and histologic criteria. Patients with indeterminate colitis were not included in this review. For each enrolled subject, clinical and demographic characteristics, including type and extent of IBD and disease activity assessment, were recorded at the time of initial diagnosis. Laboratory data, including hemoglobin level, platelet count, albumin level, and erythrocyte sedimentation rate (ESR), were also recorded on standardized forms and transmitted to a central data repository. Approval for the registry was received from the human subjects review committee at each participating institution. Informed consent was obtained from all families.

Laboratory Values

Hemoglobin levels were considered normal if they were ≥11.0 g/L for patients of either gender who were <6 years of age, ≥11.5 g/L for patients of either gender who were ≥6 and <12 years of age, ≥12 g/L for female patients who were ≥12 years of age, and ≥13 g/L for male patients who were ≥12 years of age.1 Platelet counts of ≤450 × 10⁹ platelets per L, albumin levels of ≥3.4 g/L, and ESR values of <20 mm/hour were considered normal for all ages and both genders. All laboratory testing was performed locally and not at a centralized laboratory.

Disease Activity

At the time of diagnosis, the attending physician categorized disease activity with physician global assessment (PGA). This assessment was made after the history and physical examination was completed but before the laboratory data were known; patients were classified as having quiescent, mild, moderate, or severe disease. The PGA has been used as the standard with which other, more-quantitative instruments to assess the activity of CD (eg, pediatric CD activity index) that include laboratory data have been compared. Previous studies showed excellent correlation between the PGA and other indices of disease activity.2,3 There has been no validated instrument for assessing the activity of UC in either adults or children, and usually the PGA is used.

Statistical Analyses

Data are shown as mean ± SD. Differences between the normal and abnormal laboratory results for each of the 4 tests studied were evaluated by using the χ² test and Fisher’s exact test. In addition, the χ² test was used for analysis of disease location in relation to disease severity. Differences between CD and UC for the disease severity groupings were evaluated by using t tests. The Mann-Whitney U test and 1-way analysis of variance with multiple comparisons were used for comparison of disease location with the number of abnormal laboratory tests. The McNemar test was used for correlations between normal and abnormal laboratory values. P < .05 was considered significant for all tests; in the case of multiple pairwise tests, a significance level of 0.05/k was used, where k represents the number of tests being performed, to maintain the overall error rate of .05. SPSS 12.0.1 for Windows (SPSS, Chicago, IL) was used by the Pediatric IBD Collaborative Research Group Registry statisticians (Ms Langton and Ms Lerer) to conduct the statistical analyses.

RESULTS

The characteristics of the study population are shown in Table 1. A total of 526 participants (58% male) were diagnosed as having CD (n = 392) or UC (n = 134). The age at onset of disease was 11.6 ± 3.1 years (mean ± SD). Seventy-one percent of children had moderate/severe disease at the time of diagnosis.

The frequency of normal laboratory tests for children with newly diagnosed CD or UC is shown in Table 2. Among children who presented with mild disease activity and had complete laboratory data for evaluation, normal values for all 4 laboratory tests were found for 21% of CD participants (22 of 105 patients) and 54% of UC participants (21 of 39 patients). Patients with moderate/severe IBD were more likely to have ≥1 abnormal
laboratory test, but 4% (14 of 353 patients) had normal values for all 4 laboratory tests.

When tests were examined individually, ESR was the least likely of the laboratory tests to yield normal results, with normal ESR results being found for 26% of all subjects (129 of 497 patients) (Table 3). Even for subjects with moderate/severe disease, a normal platelet count or albumin level was found for 43% and 50% of patients, respectively, whereas 24% had normal hemoglobin levels and 18% had normal ESR results. Children with mild CD had normal laboratory results more frequently than did children with either moderate or severe CD (Table 3). Similarly, patients with mild UC had normal ESR values, platelet counts, and hemoglobin levels more frequently than did those with either moderate or severe disease (Table 3). Whereas rates of normal albumin levels were similar for UC participants with mild and moderate disease, both groups were more likely to have normal albumin levels than were children with severe UC (Table 3).

The numerical values of ESRs, platelet counts, and albumin levels, are shown in Fig 1. These numerical values for mild, moderate, and severe CD and UC reflect the frequency of normal laboratory findings discussed in the previous section, with greater disease severity having median values further from the normal outer range. Mean values of ESR (P = .026) and platelet counts (P < .001) in mild CD were higher than those in mild UC. Mean values of ESR (P = .026), platelet counts (P = .021), and albumin levels (P = .001) were worse in moderate CD than in moderate UC.

Because normal hemoglobin levels vary with age and gender, comparisons were made for patients 6 to 12 years of age and ≥12 years of age, to minimize the age effect. For patients 6 to 12 years of age, the hemoglobin level in mild CD (12.1 ± 4.0 g/L; mean ± SD; n = 47) was similar to the hemoglobin level in mild UC (12.2 ± 1.3 g/L; n = 19), and the hemoglobin level in moderate CD (11.2 ± 2.4 g/L; n = 69) was similar to that in moderate UC (10.4 ± 1.8 g/L; n = 28). However, the hemoglobin level in severe CD (10.7 ± 1.5 g/L; n = 30) was higher than that in severe UC (8.3 ± 0.7 g/L; n = 4; P = .006). Similarly, for patients ≥12 years of age, the hemoglobin levels in mild CD (12.6 ± 3.4 g/L; n = 49) and mild UC (12.5 ± 1.5 g/L; n = 15) were similar. Hemoglobin levels in CD (11.3 ± 1.5 g/L; n = 12) and moderate UC (11.6 ± 1.6 g/L; n = 35) were similar to each other. Unlike in the younger age group, however, hemoglobin levels in severe CD (10.5 ± 1.8 g/L; n = 33) and severe UC (9.4 ± 1.8 g/L; n = 15) were similar.

As shown in Table 1, CD was more likely to involve multiple regions of the gastrointestinal tract than a single location (ie, upper gastrointestinal tract, small bowel, or colon). For UC, involvement of the entire colon was the most likely distribution of disease at the time of diagnosis. Figure 2A reveals no association between location of disease and severity of disease in CD. There was no clear pattern for the number of normal laboratory values for participants with CD (Fig 3A). Pairwise analysis between the different locations of bowel affected by CD revealed colon-only disease was more likely to have normal tests than small bowel and ascending colon disease (P = .034) or disease involving the upper gastrointestinal tract and small bowel with or without any region of the colon (P < .001). However, these findings do not take into account the length of mucosa involved in each of these regions. As might be expected, there was a direct correlation between disease severity and extent of bowel involved in UC (Fig 2B) and the number of normal labo-

![Table 1](https://example.com/table1.png) **Characteristics of the Study Population**

<table>
<thead>
<tr>
<th>CD</th>
<th>No. of subjects</th>
<th>Age at diagnosis, mean ± SD, y</th>
<th>PGA, %</th>
<th>Location, %</th>
<th>Moderate</th>
<th>Severe</th>
<th>Male/female, %</th>
<th>White/nonwhite race, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of subjects</td>
<td>392</td>
<td>11.8 ± 2.9</td>
<td>29</td>
<td>1</td>
<td>53</td>
<td>18</td>
<td>59/41</td>
<td>86/14</td>
</tr>
<tr>
<td>White/nonwhite race, %</td>
<td>100/100</td>
<td></td>
<td></td>
<td>Upper gastrointestinal tract only</td>
<td>1</td>
<td>8</td>
<td>Small bowel only</td>
<td>28/28</td>
</tr>
<tr>
<td>Moderate</td>
<td>134</td>
<td>11.2 ± 3.6</td>
<td>30</td>
<td>Small bowel and colon</td>
<td>55</td>
<td>15</td>
<td>75/25</td>
<td>85/15</td>
</tr>
<tr>
<td>Severe</td>
<td>134</td>
<td>11.2 ± 3.6</td>
<td>55</td>
<td>Colon only</td>
<td>13</td>
<td>5</td>
<td>45/55</td>
<td>65/35</td>
</tr>
<tr>
<td>Male/female, %</td>
<td>60/40</td>
<td></td>
<td></td>
<td>Upper bowel, small bowel, and/or colon</td>
<td>77</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Upper gastrointestinal tract includes esophagus and/or stomach.

![Table 2](https://example.com/table2.png) **Frequency of Normal Laboratory Values at Diagnosis**

<table>
<thead>
<tr>
<th>Disease Subtype</th>
<th>Frequency of Normal Laboratory Values, %</th>
<th>0</th>
<th>Tests 4</th>
<th>Tests 3</th>
<th>Tests 2</th>
<th>Tests 1</th>
<th>Tests 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>21</td>
<td>25</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>5</td>
<td>34</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>2</td>
<td>32</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
<td>31</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>54</td>
<td>8</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>6</td>
<td>26</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>0</td>
<td>45</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>24</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
disciplinary tests (Fig 3B), with pancolitis being worse than either left-sided disease or isolated rectosigmoid disease. Taken together, our results showed that, of patients with only rectosigmoid involvement, 71% had mild disease and 79% had 4 normal laboratory values. In contrast, of patients with pancolitis, 79% were considered to have moderate/severe disease and only 14% (14 of 103 patients) had normal values for all 4 laboratory tests.

Overall, 41 of 144 children with mild IBD and normal values for all 4 of the laboratory tests evaluated had all clinical information available at the time of diagnosis. The 3 most common presenting features among the subjects with CD were blood in the stools (65%; 13 of 20 patients), abdominal pain (60%; 12 of 20 patients), and complications. For some children eventually diagnosed as having IBD, the time from onset of symptoms to diagnosis of IBD can be long.4,5 Our data did not identify what factors in the presentation of our patients might have prompted additional diagnostic studies for IBD, even in the presence of normal screening laboratory studies. However, many of the children diagnosed as having mild IBD had reports of blood in the stools and possibly other nonspecific complaints (eg, abdominal pain, poor energy, or diarrhea). Therefore, it is clear that physicians must remain suspicious regarding the possibility of IBD in patients with symptoms of intestinal inflammation, irrespective of laboratory findings.

DISCUSSION
The finding that a significant number of children with IBD have a normal battery of screening laboratory tests at the time of diagnosis has clear implications for primary health care providers. For a child with a history of mild abdominal pain or diarrhea, the presence of normal screening laboratory studies often implicates functional disorders (eg, irritable bowel syndrome) as a cause of symptoms and precludes additional diagnostic testing. Although functional disorders are much more common than IBD, our data show that the presence of normal laboratory test results cannot be relied on as an adequate screening tool to exclude mild IBD. Active IBD is associated with a number of clinical signs and symptoms; if these are present, then additional investigation is warranted. In our study, the most common sign for those with mild IBD was hematochezia, and investigation for children with this problem is warranted even if the laboratory test results are normal. In contrast, children with more-severe IBD only rarely have all 4 of the laboratory tests yielding normal results at presentation.

Failure to diagnose IBD in children with mild symptoms can lead to delayed diagnosis, which can result in more-active disease associated with increased morbidity and complications. For some children eventually diagnosed as having IBD, the time from onset of symptoms to diagnosis of IBD may be long.6,7 Although disease severity was not indicated in the study by Beattie et al,6 it seems that the patients in the moderate/severe category, not only on the
basis of the elevated CRP levels but also because the likelihood of an elevated ESR of 85% for their patients was similar to what we found for the patients in the moderate/severe CD category in the current study. Therefore, we conclude that, even if all of the simple blood tests yield normal results, children should be considered for referral for additional evaluation when there are chronic gastrointestinal complaints, such as hema-

tochezia.

There have been other studies that have evaluated laboratory abnormalities for children and adolescents with IBD. The sample sizes in those previous studies were also small, ranging between 24 and 100 subjects; in addition, many of those reports did not include variables such as disease severity and/or extent of disease involve-

FIGURE 1
Individual laboratory values in new-onset IBD. Box plots are shown for the distribution of the numerical values of ESR (A), albumin level (B), and platelet count (C) for participants with CD and UC at the time of diagnosis. Box plots show median values within each box (25th and 75th percentiles at the bottom and top of each box, respectively), error bars extending from each box are the 5th and 95th percentiles, and circles indicate values outside these representations.

FIGURE 2
Disease severity assessed according to region of disease involvement. Values are frequencies, expressed as percentages of CD (A) and UC (B) participants with mild (light gray bar), moderate (dark gray bar), or severe (black bar) disease for the different sites of bowel involvement. GI indicates gastrointestinal tract.
ment. In a systematic review of anemia that included pediatric and adult studies, the general prevalence of anemia in IBD ranged between 10% and 73% in CD and between 9% and 67% in UC. Worse disease severity and younger age were correlated with the frequency of anemia. At diagnosis in our study, we found anemia to be common, with 69% of children with CD (252 of 366 children) and 64% of children with UC (84 of 131 children) being affected. Similar to our data, worse disease severity was correlated directly with frequency of anemia. Weinstein et al reviewed laboratory test results for children with newly diagnosed CD (n = 82) or UC (n = 71). They reported that, at the time of presentation, absolute laboratory values for ESR and platelet count were higher in CD than in UC and absolute values for hemoglobin and albumin levels were lower in CD than in UC. With adjustment for disease severity, we could not confirm this finding for severe disease, because at presentation the laboratory values were similar for CD and UC. Furthermore, even at mild and moderate disease severity, there is considerable overlap between laboratory values and clinically it is unlikely that this would be a useful parameter to differentiate subtypes of IBD.

Additional laboratory markers to help diagnose and differentiate subtypes of IBD and to assess IBD disease severity and extent of disease could help in reducing risks and costs to patients, because current evaluation strategies involve labor-intensive, potentially risky, costly methods, including radiologic studies and endoscopic studies with histologic evaluation of specimens, in addition to traditional laboratory testing. Newer and more costly laboratory evaluations have been suggested as ways to discriminate between IBD and other causes of chronic gastrointestinal symptoms, such as irritable bowel syndrome, and to define an individual’s IBD. With inflammation of the lining of the gastrointestinal tract being the primary source of inflammation in IBD, detection of a variety of stool markers mostly derived from leukocytes (e.g., lactoferrin, elastase, lysozyme, myeloperoxidase, calprotectin, and S100 proteins) has been studied. The noninvasive nature of fecal tests makes them ideal candidates for laboratory markers, but lack of extensive analysis, availability, and costs are drawbacks. Although these tests can have high specificity for inflammation of the intestinal tract, because their source is from leukocytes, they are not specific for IBD, and detection levels may not be equal among the subtypes of IBD.

Another route for evaluation of laboratory markers has been the evaluation of specific serologic markers for patients with IBD, which has yielded a number of antibodies, including antibodies against neutrophils (atypical, perinuclear, cytoplasmic, DNase-sensitive, antineutrophil antibodies), antibodies against microbial antigens (anti-\textit{Saccharomyces cerevisiae} antibodies, anti-outer membrane porin C antibodies, anti-I2 antibodies, and anti-flagellin antibodies), and antiglycan antibodies. The combining of >1 of the markers has produced excellent specificity and good sensitivity and may facilitate diagnosis for some patients, but for most patients the diagnosis is suspected without marker use. These serologic markers may also define subsets of patients with IBD, but, for initial diagnostic tests, it should be kept in mind that they are not available in many individual testing facilities, they are expensive, they do not measure disease activity, and they do not determine the site and extent of disease, the latter of which are used to determine management strategies at the current time. Moreover, for young patients with IBD, it is less likely that serologic tests would be positive, which raises the possibility that these antibodies may be secondary phenom-
enon, may take time to develop, or may not be definable in this subgroup of patients.

Use of a strategy involving only traditional laboratory tests to screen for possible IBD would miss significant numbers of children with primarily mild IBD. However, good clinicians rely on more than just laboratory assessments when making judgments about patients. In particular, the presence of blood in the stools should be a warning sign that a child with normal laboratory results requires additional evaluation. Although clinicians should be reassured that 94% to 98% of children with IBD would be identified with \( \geq 1 \) abnormal laboratory test or the presence of blood in the stools, clinical suspicion remains critical in the decision-making process, to direct additional diagnostic testing.

**ACKNOWLEDGMENTS**

Financial support for this study was provided by the Ottawa Snowflake Ball, Centocor (Malvern, PA), Reach Out for Youth with Ileitis and Colitis (Melville, NY), and the collaborating institutions.

We are deeply indebted to the following research coordinators, whose efforts greatly facilitated the performance of this study: Ruth Singleton, Patricia Davis, Kathy Grancher, Valerie Grant, Annette Langseder, Anna Zholudev, George Kay, Gail Waltz, Kim Boyer, Shari Huffman, Cathy Williams, Rebecca Abood, Rosemary Nagy, Carol Rudman, Myrna Miller, Vivian Abadom, Janet Trotta, and Laura Defaveri.

**REFERENCES**

Role of Zinc Administration in Prevention of Childhood Diarrhea and Respiratory Illnesses: A Meta-analysis

Rakesh Aggarwal, MD, DMa,b, John Sentz, MPHc, Mark A. Miller, MDc

aDepartment of Gastroenterology, Sanjay Gandhi Postgraduate Institute of Medical Sciences, Lucknow, India; bDepartment of Gastroenterology, All India Institute of Medical Sciences, New Delhi, India; cDivision of International Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health, Bethesda, Maryland

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

ABSTRACT

BACKGROUND. The quantified effect of zinc supplementation to prevent childhood diarrhea and respiratory illnesses is unclear. We conducted a meta-analysis of randomized, controlled trials on the subject.

METHODS. We searched PubMed, Science Citation Index, and the Cochrane Database of Controlled Trials and hand-searched the reference lists of identified articles. All randomized, controlled trials of zinc supplementation for ≥3 months for children <5 years of age, using blinded assessment, were eligible. The outcome measures studied were number of episodes of illness, number of days with illness, and number of episodes of severe illness. Data from 17 studies were pooled by using random-effects and fixed-effects models for data with and without significant heterogeneity, respectively.

RESULTS. Children who received a zinc supplement had fewer episodes of diarrhea (rate ratio: 0.86) and respiratory tract infections (rate ratio: 0.92) and significantly fewer attacks of severe diarrhea or dysentery (rate ratio: 0.85), persistent diarrhea (rate ratio: 0.75), and lower respiratory tract infection or pneumonia (rate ratio: 0.80) than did those who received placebo. They also had significantly fewer total days with diarrhea (rate ratio: 0.86) but not days with respiratory illness (rate ratio: 0.95). Published studies showed a publication bias and significant heterogeneity; however, no cause for the latter could be identified.

CONCLUSIONS. Zinc supplementation reduced significantly the frequency and severity of diarrhea and respiratory illnesses and the duration of diarrheal morbidity. The relatively limited reduction in morbidity and the presence of significant heterogeneity and of publication bias indicate the need for larger, high-quality studies to identify subpopulations most likely to benefit.
Zinc is a vital micronutrient in humans and is essential for protein synthesis, cell growth, and differentiation.4,2 Severe zinc deficiency has been shown to be associated with stunting of growth, hypogonadism, impaired immune function, skin disorders, cognitive dysfunction, and anorexia.2,3

Dietary deficiency of zinc is common in several parts of the world, particularly developing countries. This deficiency may arise either from inadequate intake of foods that contain zinc (mainly foods of animal origin) or from inadequate absorption caused by its binding to dietary fiber and phytates, which are commonly found in cereals, nuts, and legumes.4 Inadequate absorption of zinc may also result from mucosal abnormalities and compromised gut integrity induced by persistent diarrhea attributable to gastrointestinal infections.5

Numerous studies have examined the association between childhood morbidity and zinc deficiency.1,6–9 Those studies suggest that zinc-deficient populations are at increased risk of developing diarrheal diseases, respiratory tract infections, and growth retardation. Using the available data, Caulfield and Black10 estimated that zinc deficiency is associated with nearly 800 000 excess deaths annually among children <5 years of age throughout the world, including deaths attributable to diarrhea (176 000), pneumonia (406 000), and malaria (207 000). Furthermore, they attributed a global loss of nearly 28 million disability-adjusted life years to zinc deficiency.

Zinc administration has been studied as a tool for the treatment and prevention of diarrhea, respiratory tract infections, pneumonia, acute lower respiratory tract infections, and malaria among children.11–18 A meta-analysis showed that addition of zinc to the treatment regimen for children with diarrhea led to reductions in the duration of diarrhea and in the frequency of persistent diarrhea, defined as diarrhea lasting >14 days.19

Those studies led to the use of zinc supplementation among children in attempts to treat and to prevent common childhood infections. A few studies also examined the effects of zinc supplementation on growth parameters.20–25 A meta-analysis published in 1999 showed that continuous zinc supplementation was associated with decreased rates of childhood diarrhea and pneumonia.26 Since then, additional studies that were larger in size and scope than those included in the meta-analysis have been published. Furthermore, the previous meta-analysis included studies with supplementation for periods as short as 2 weeks. Supplementation programs are likely to be provided to children for longer periods (at least a few months). Therefore, we thought that the subject merited reexamination, and we conducted a meta-analysis of studies that examined the efficacy of zinc supplementation lasting ≥3 months in preventing diarrhea and respiratory illnesses among children.

METHODS

Search Protocol and Study Review
To identify studies that examined the effect of zinc supplementation on the occurrence of diarrhea or respiratory tract infections, PubMed, Science Citation Index, and the Cochrane Central Database of Controlled Trials were searched by using the following keywords: “zinc,” “supplement,” and “diarrhea” or “respiratory illness” or “pneumonia.” The searches captured studies published up to November 2005. Two independent reviewers reviewed the search results to identify relevant original human clinical or field trials. Studies that focused on the effects of zinc administration for the treatment of acute or persistent diarrhea or respiratory illnesses were excluded from the analysis. However, studies that enrolled children who had recently recovered from a diarrheal illness and were being observed for subsequent recurrent diarrheal episodes were deemed eligible for inclusion in the meta-analysis. Studies with zinc supplementation for periods of <3 months were excluded. Additional studies were identified through manual searches of reference lists of the originally identified studies on the therapeutic and preventive roles of zinc, as well as reviews on the subject.

All identified studies were reviewed independently by 2 authors (Dr Aggarwal and Mr Sentz), to determine whether they fulfilled the minimal quality criteria, including (1) random allocation of placebo and active interventions, (2) double-blinded assessment of outcomes, and (3) ≥90% follow-up rates. Furthermore, the reviewers identified the parameters for which data were available in the reports, for use in designing a data extraction form. The reviewers extracted data from each of the selected studies independently; any inter-reviewer differences were resolved through a joint review of the article.

Statistical Methods
From the selected studies, data on the number of episodes of diarrhea and respiratory illnesses in the groups receiving zinc supplementation and placebo were extracted. When such data were not provided explicitly, they were calculated by multiplying the mean numbers of episodes of these illnesses per child by the total number of study subjects in each group. From the number of disease episodes and the duration of follow-up monitoring for the supplementation and placebo groups, adjusted rate ratios (RRs) were calculated for each study. In addition, data on rates of severe forms of diseases of interest (eg, persistent diarrhea lasting ≥14 days, severe diarrhea [based on investigator-defined criteria], lower respiratory tract infection, or pneumonia) were retrieved when available.

Statistical analyses and meta-analyses were performed by using Review Manager 4.2.8 software (Nordic
Cochrane Centre, Copenhagen, Denmark). SEs for logarithms of RRs were calculated by using standard statistical methods. Data from different studies were pooled by using a generic inverse variance method, and a pooled RR, SE, and 95% confidence interval (CI) were calculated for each parameter. For each parameter, heterogeneity between studies was tested by using the $\chi^2$ statistic with its degrees of freedom; in addition, the $I^2$ statistic, measuring the extent of inconsistency of results between various studies, was calculated. Significant heterogeneity was considered to be present if the $P$ value was below .10. Where there was no significant heterogeneity ($P \geq .10$), a fixed-effects model was used for pooling of data from various studies; in cases with significant heterogeneity ($P < .10$), a random-effects model was used to provide a more-conservative estimate of effect. In the latter cases, the data were also analyzed by using subgroup analyses, to assess possible causes of heterogeneity.

Publication bias was examined by using “funnel plot” analysis and was quantified by using the rank correlation method described by Begg and Mazumdar27 and the regression intercept method described by Egger et al,28 with $P$ values of <.10 being taken as significant. In addition, the trim and fill method described by Duval and Tweedie29,30 was used to impute missing studies and to recompute the combined effect by adding those missing studies.

RESULTS

Identification of Studies

Table 1 lists the clinical or field trials that were identified in our literature search as having evaluated the role of zinc in the prevention of diarrhea or respiratory illnesses.15,16,23,24,31–43 Four additional studies21,22,25,44 that examined the preventive effects of zinc supplementation on these illnesses were excluded for various reasons. Of those, one study was excluded because the assignment of study subjects to zinc or placebo groups was nonrandom.23 One study was considered ineligible because it studied morbidity outcomes among newborn infants whose mothers had received zinc supplements during pregnancy.44 A study conducted in Brazil included 2 separate zinc supplementation groups (1 and 5 mg/day); allocation to the latter group was neither random nor concurrent with that to the placebo group, making the group ineligible for the meta-analysis.31 Data from the other zinc supplementation arm of the study (1 mg/day) had to be excluded because the published data did not permit extraction of the parameters that were used in our analysis. Finally, in one study,25 zinc and “psychosocial stimulation,” an intervention aimed at improving mother-child interaction, were used in a factorial study design, with allocations to the interventions performed in a random manner and a nonrandom manner, respec-
tively. That study had to be excluded because the published report did not allow extraction of data for groups that received zinc or placebo without psychosocial stimulation. Also, the data were presented as group medians, which precluded calculation of total numbers of illness episodes.

A total of 17 studies15,16,23,24,31–42 were included in the meta-analysis. Of those, one study36 had 2 intervention arms, with differing frequencies (daily or weekly) of zinc supplement administration; for analysis, each of the treatment arms was considered a separate study. The subsequent text refers to those as independent studies. Two published studies reported on the same population separately for diarrheal and respiratory morbidity outcomes.35,43 Data from the 17 studies were analyzed and included 3819 children who received zinc supplementation and 3840 children who received placebo. Table 2 lists some of the other excluded studies, with reasons for their exclusion.20,45–48

Prevention of Diarrhea
Of the 17 eligible studies, 15 provided data on the number of episodes of diarrhea among groups that received a zinc supplement or a matched placebo.15,16,23,24,31–40 Figure 1 shows the results of the meta-analysis for the incidence of diarrhea among children who received zinc supplements or a matched placebo. Of the 15 studies, 3 had RRs of >1.0 and 12 RRs of <1.0. With a random-effects model, our analysis revealed that zinc supplementation was associated with a significant reduction in the occurrence of diarrheal episodes, by 14% (RR: 0.86; 95% CI: 0.79–0.93).

Prevention of Respiratory Illness
Data on the frequency of respiratory illnesses were available in 12 studies, which included 2709 children who received a zinc supplement and 2803 children who received a matched placebo.15,16,23,24,31–34,37,39,40,43 The pooled data showed an 8% reduction in the occurrence of respiratory illness among children who received a zinc supplement, with a random-effects, pooled RR of 0.92 (95% CI: 0.85–0.99) (Fig 2).

Prevention of Severe Forms of Diarrheal Illness
Data on RRs for the incidence of severe diarrheal disease, including severe diarrhea or dysenteric illness, were available in 5 studies.15,31,33,36 The fixed-effect pooled RR

<table>
<thead>
<tr>
<th>Study or Subcategory</th>
<th>RR (Random) 95% CI</th>
<th>Weight, %</th>
<th>RR (Random) 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gupta et al, 2003 (daily)</td>
<td>1.82</td>
<td>0.41 [0.24–0.71]</td>
<td></td>
</tr>
<tr>
<td>Gupta et al, 2003 (weekly)</td>
<td>1.82</td>
<td>0.41 [0.24–0.71]</td>
<td></td>
</tr>
<tr>
<td>Umez et al, 2000</td>
<td>2.47</td>
<td>0.46 [0.29–0.72]</td>
<td></td>
</tr>
<tr>
<td>Rosado et al, 1997</td>
<td>3.05</td>
<td>0.63 [0.42–0.94]</td>
<td></td>
</tr>
<tr>
<td>Ninh et al, 1996</td>
<td>3.82</td>
<td>0.54 [0.38–0.75]</td>
<td></td>
</tr>
<tr>
<td>Bates et al, 1993</td>
<td>4.12</td>
<td>1.24 [0.90–1.71]</td>
<td></td>
</tr>
<tr>
<td>Sur et al, 2003</td>
<td>4.17</td>
<td>0.71 [0.51–0.97]</td>
<td></td>
</tr>
<tr>
<td>Osendarp et al, 2002</td>
<td>7.47</td>
<td>1.11 [0.93–1.34]</td>
<td></td>
</tr>
<tr>
<td>Lind et al, 2004</td>
<td>7.55</td>
<td>1.07 [0.89–1.28]</td>
<td></td>
</tr>
<tr>
<td>Baqui et al, 2003</td>
<td>9.87</td>
<td>0.98 [0.87–1.09]</td>
<td></td>
</tr>
<tr>
<td>Penny et al, 2004</td>
<td>10.03</td>
<td>0.89 [0.80–0.99]</td>
<td></td>
</tr>
<tr>
<td>Ruel et al, 1997</td>
<td>10.03</td>
<td>0.78 [0.70–0.87]</td>
<td></td>
</tr>
<tr>
<td>Sazawal et al, 1997</td>
<td>10.66</td>
<td>0.92 [0.84–1.00]</td>
<td></td>
</tr>
<tr>
<td>Brooks et al, 2005</td>
<td>11.46</td>
<td>0.94 [0.89–0.99]</td>
<td></td>
</tr>
<tr>
<td>Bhaddar et al, 2002</td>
<td>11.67</td>
<td>0.90 [0.85–0.95]</td>
<td></td>
</tr>
</tbody>
</table>

Total (95% CI)
Test for heterogeneity: χ² = 61.79, df = 14 (P < 0.0001), P = 77.3% Test for overall effect: Z = 3.69 (P = 0.0002)
for such illness in the children who received a zinc supplement, compared with the children who received placebo, was 0.85 (95% CI: 0.75–0.95), which indicated a significant reduction in the frequency of severe diarrhea (Fig 3). With the random-effects model, the pooled RR was 0.84 (95% CI: 0.70–1.01).

Three studies provided data on RRs for the incidence of episodes of persistent diarrhea. A pooled analysis of those studies using a fixed-effect model showed a significant reduction in such illness, with a pooled RR of 0.75 (95% CI: 0.57–0.98) for the occurrence of persistent diarrhea in the children who received a zinc supplement, compared with the children who received a placebo (Fig 4). Use of the random-effects model was not possible because of the small number of available studies.

**Prevention of Severe Respiratory Illness**

Four studies provided data for comparison of incidence rates of severe respiratory illness, described variably as pneumonia or lower respiratory tract infection. The fixed-effects pooled estimate of the RR for such illness was 0.80 (95% CI: 0.70–0.92) (Fig 5), which indicated a significant reduction in the frequency of such illness. With the random-effects model, the pooled RR was 0.79 (95% CI: 0.65–0.95).

**Duration of Diarrhea and Respiratory Illnesses**

The pooled efficacies of zinc in the reduction of the number of days with diarrhea or respiratory illnesses were 0.86 (95% CI: 0.79–0.93) and 0.95 (95% CI: 0.84–1.07), respectively (Figs 6 and 7); the latter was not statistically significant.

---

**FIGURE 2**

Meta-analysis of RRs of incidence of respiratory illness episodes in children who received zinc supplementation or a placebo.

**FIGURE 3**

Meta-analysis of RRs of incidence of episodes of severe diarrhea and/or dysentery in children who received zinc supplementation or a placebo.
Reasons for Heterogeneity Among Studies

Subgroup analyses for the type of zinc compound used, the total dosage of zinc per week ($\geq 70$ mg versus $< 70$ mg), and the frequency of zinc administration (daily or less frequently) did not explain the heterogeneity between studies. The available data did not permit evaluation of the effects of zinc supplementation separately for children with good versus poor nutrition or children with zinc deficiency versus no zinc deficiency at the time of inclusion in the study.
Publication Bias
To assess whether there was a bias in the published literature toward studies with positive outcomes, we plotted the effect size of each trial versus variance of the effect for reduction in the number of episodes of diarrhea or respiratory illnesses. In the absence of a publication bias, such a plot is expected to have a shape resembling an inverted funnel. The funnel plots for both parameters were nearly identical (Fig 8), with an excess of favorable studies with high variance and an apparently limited number of published studies of small size and negative effect, which indicated the existence of a publication bias.

The regression intercept method described by Egger et al indicated the presence of significant publication bias in studies on the effects of zinc supplementation on diarrhea ($P = 0.070$) but not respiratory illnesses ($P = 0.340$). The rank correlation method described by Begg and Mazumdar failed to show the presence of significant publication bias in studies on diarrhea ($P = 0.138$) but found evidence of such bias in studies on respiratory illnesses ($P = 0.029$). The trim and fill analysis for studies on diarrhea showed 3 missing studies in the right lower portion; with the addition of those studies, a summary RR of 0.90 (95% CI: 0.82–0.98) was obtained. Similarly, there were 3 missing studies for respiratory illnesses; with the addition of those studies, the summary RR was estimated as 0.95 (95% CI: 0.87–1.04).

DISCUSSION
Our meta-analysis indicated that zinc supplementation for young children led to reductions in the risk of diarrhea and respiratory tract infections (14% and 8%, respectively). Zinc supplementation was also associated with reductions in the rates of serious forms of these illnesses and in the number of days of diarrhea per child. However, there was no significant reduction in the number of days with respiratory illness. The published literature showed significant heterogeneity regarding the preventive effects of zinc administration, with evidence of publication bias.

The current meta-analysis focused on estimating the

![Figure 7](image1.png)

*FIGURE 7*  
Meta-analysis of RRs of number of days with respiratory illness in children who received zinc supplementation or a placebo. OR indicates odds ratio.

![Figure 8](image2.png)

*FIGURE 8*  
Funnel plot of studies on zinc supplementation for prevention of diarrhea (A) and respiratory illnesses (B) to assess the presence of publication bias.
effect of zinc supplementation on the reduction of childhood morbidity. Zinc deficiency, arising from inadequate dietary intake or poor absorption, is common in many developing countries. It is thought to be 1 of the 10 greatest factors contributing to disease burden among children in developing countries and has led to calls for the initiation of supplementation and food fortification programs. Although several studies have assessed the efficacy of zinc supplementation in preventing infections among children, the results of those studies were quite variable. The current meta-analysis provides a structured review to summarize the effects of zinc supplementation and provides quantified effectiveness data, to aid in policy formulation regarding the implementation of large-scale zinc supplementation programs. It provides evidence that zinc supplementation programs are likely to reduce morbidity attributable to diarrhea and respiratory tract infections, although the proportions prevented are likely to be small.

Our findings of reduced frequency of diarrhea and respiratory illnesses in children receiving zinc supplementation indicate that this intervention could be useful in developing countries where zinc deficiency is common and mortality rates are high. Although the effects of zinc supplementation on the morbidity attributable to childhood infections were modest, the absolute number of illness episodes prevented would be large, given that most children <5 years of age suffer several episodes per year, on average. Given that >3.5 million children are estimated to die as a result of either diarrhea or respiratory illnesses, even small effects could translate into major absolute reductions in childhood morbidity and mortality rates.

The estimates provided by the current meta-analysis should allow for better estimation of the benefits of zinc supplementation. Throughout the world, nearly 2 billion episodes of diarrhea are estimated to occur every year among children <5 years of age. With the pooled estimate from our study, zinc supplementation might be expected to prevent 280 million episodes of diarrhea each year. Furthermore, the data from this meta-analysis should be useful for calculation of cost-effectiveness and cost/benefit ratios for administration of zinc supplements in developing countries.

The relatively modest effect of zinc supplementation observed in the current meta-analysis implies that this intervention should be combined with other interventions aimed at reducing childhood morbidity. The relative financial feasibility and cost-effectiveness of such interventions could help determine their priority for implementation. Alternatively, a combination of interventions may be synergistic, with the combination providing greater benefits than the sum of benefits expected from the measures applied individually, thus enhancing their cost-effectiveness.

Our meta-analysis has several strengths. First, it included only studies that were deemed high quality by meeting strict inclusion criteria. It included only randomized, placebo-controlled studies in which assessors were unaware of treatment allocation. Blinded assessment prevents observer bias; this may be particularly important in studies pertaining to zinc supplementation, because major outcome measurements in such studies (i.e., occurrence of diarrhea or respiratory illnesses and their severe forms) are subjective in nature. Second, we took care to exclude studies in which zinc administration was performed with a therapeutic intent, whereas a previous meta-analysis on this subject included such studies, which might have led to an overestimation of the potential preventative effects of zinc. Furthermore, our meta-analysis included the results of larger studies on the subject. In addition, we used a more-robust, random-effects model for pooling the results of published studies if their results had significant heterogeneity.

Our analysis has some inherent limitations. It did not take into consideration the beneficial effect of zinc supplementation on linear growth, which is often cited as another benefit of zinc supplementation. We also did not consider the effect of zinc supplementation on malaria. The effect of zinc supplementation on malaria has been examined in only a few studies in Africa, which were too limited, compared with those for diarrhea and respiratory illnesses. Also, our results may not be applicable to children with HIV infection, although diarrhea and respiratory illnesses are common in such children. Furthermore, our meta-analysis addressed only prophylactic effects of zinc and did not review studies on the therapeutic effects of zinc.

The benefits of any supplementation program, including those for zinc administration, may be expected to be disproportionately larger for subjects with marginal or poor nutritional status. In 2 studies, children who were enrolled with initially low serum zinc levels seemed to experience greater reductions in the incidence and prevalence rates of diarrhea after receiving zinc, compared with children with higher baseline serum zinc levels. However, it was not possible to analyze data on diarrhea and respiratory illnesses for children with varying degrees of malnutrition, because such stratified data were not available in the other studies that met our inclusion criteria. In fact, in trials in which zinc supplementation was tried for the treatment of diarrhea, it seemed to have greater value for children with poor nutritional status and serum zinc levels. Future studies that could analyze the different effects of zinc on children who are deemed zinc deficient or not would be useful for identifying subpopulations that could most benefit in resource-limited settings. It should be noted that the diagnosis of zinc deficiency remains difficult, and serum zinc levels are not necessarily accurate for this purpose. Furthermore, most of the zinc supplemen-
tation studies were performed in developing countries, and the results of our meta-analysis may not be applicable to children in developed regions. Several large, randomized, controlled studies undertaken after the 1999 meta-analysis supported the results presented in the current meta-analysis.16,35,40

The existence of significant heterogeneity in the results of various published studies may distract from the conclusions of our meta-analysis. We accounted for this by using random-effects, meta-analysis techniques for analyses in which studies showed significant heterogeneity. The fixed-effects model was used only for some analyses; in those analyses, we also pooled data by using the random-effects model, and we found the results to be largely similar.

The funnel plot analysis indicated the presence of publication bias. This visual impression was supported by quantitative measures of publication bias; the discordance between the results of the methods for measuring publication bias described by Begg and Mazumdar27 and by Egger et al28 can be explained by the relatively low sensitivity of these methods, whereby nonsignificant results with these tests do not rule out the presence of bias. The presence of heterogeneity and publication bias suggests the need for additional large, randomized, controlled studies to examine the benefits of zinc supplementation. It will also be important to monitor the populations for which prophylactic zinc supplementation has already been introduced, to study trends in disease morbidity rates and to estimate the effectiveness of this intervention under field conditions.

The studies included in our meta-analysis used zinc doses of 15 to 140 mg/week. Although our subgroup analyses failed to show any effect of dose on the benefits seen with zinc supplementation, this might have been related to the limited sensitivity of such analyses and the small number of studies using various doses. The large variation in the doses used in various studies, at times exceeding the recommended daily allowance for zinc (11 mg/day for men and 8 mg/day for women), suggests the need for additional studies on optimal supplementation doses, particular because high zinc intakes have been shown to be associated with inhibition of absorption of other micronutrients51 and with poorer survival rates for children with HIV infection.52

CONCLUSIONS

Pooled data from our meta-analysis indicated that zinc supplementation in healthy children led to significant but modest reductions in the frequency of diarrhea and respiratory illnesses. This intervention also led to reductions in the frequency of severe diarrhea and lower respiratory tract infections and in the number of days with diarrhea per child; however, there was no significant reduction in the number of days with respiratory illness per child in the pooled data from the included studies. These data may have public health significance, and they provide support for the implementation of zinc supplementation programs in developing countries in an attempt to improve child health. Furthermore, the results of our analysis emphasize the need for additional data collection pertaining to zinc in larger studies, especially those in which study subjects are stratified according to baseline nutritional status. Such data would be important in evaluating differences in the responses of well-nourished and poorly nourished children to zinc supplements and might thus enable better targeting of this intervention.

ACKNOWLEDGMENTS

Dr Aggarwal was supported by the Overseas Associateship Program of the Department of Biotechnology, Government of India, during this work. This work was partly funded by the Fogarty International Center, National Institutes of Health, and the Bill and Melinda Gates Foundation.

We thank Dr Robert Black and Jessica Seidman for critical review of the manuscript and suggestions.

REFERENCES
13. Roy SK, Tomkins AM, Mahalanabis D, et al. Impact of zinc...


47. Baqui AH, Black RE, EL Arifeen S, et al. Effect of zinc supplementation started during diarrhoea on morbidity and mortality
in Bangladeshi children: community randomised trial. BMJ. 2002;325:1059
52. Tang AM, Graham NM, Saah AJ. Effects of micronutrient intake on survival in human immunodeficiency virus type 1 infection. Am J Epidemiol. 1996;143:1244–1256

PATIENTS’ VIEWS ON FINANCIAL CONFLICTS OF INTEREST IN CANCER RESEARCH TRIALS

Background: Financial ties between researchers or medical centers and companies whose drugs are being tested have come under increasing scrutiny.

Methods: We conducted in-person interviews with 253 patients in cancer-research trials (a 93% response rate) at five US medical centers to determine their attitudes regarding potential financial conflicts of interest among researchers and medical centers.

Results: More than 90% of patients expressed little or no worry about financial ties that researchers or institutions might have with drug companies. Most patients said they would have enrolled in the trial even if the drug company had paid the researcher for speaking (82% of those interviewed) or consulting (75%) or if the researcher had received royalty payments (70%) or owned stock in the company (76%). Similarly, most patients would have enrolled in the trial if their cancer center had owned stock in the drug company (77%) or received royalty payments from the company (79%). Most patients believed it was ethical for researchers to receive speaking fees (81%) or consulting fees (82%) from the company. However, a substantial minority of patients wanted disclosure of the oversight system for researchers (40%) and of researchers’ financial interests (31%); 17% thought no disclosure to patients was necessary.

Conclusions: Most patients in cancer-research trials were not worried about financial ties between researchers or medical centers and drug companies and would still have enrolled in the trial if they had known about such financial ties. A substantial minority wanted to be informed about the oversight system to protect against financial conflicts of interest and about researchers’ financial interests.


Noted by JFL, MD
Quality of Primary Care and Subsequent Pediatric Emergency Department Utilization

David C. Brousseau, MD, MSa,b, Raymond G. Hoffmann, PhDb, Ann B. Nattinger, MD, MPHb,c, Glenn Flores, MDa,b,d, Yinghua Zhang, MSNb, Marc Gorelick, MD, MSCEa,b

aDepartment of Pediatrics, Children’s Research Institute, bDepartment of Population Health, cDepartment of Medicine, Patient Care and Outcomes Research Center, and dCenter for the Advancement of Underserved Children, Medical College of Wisconsin and Children’s Hospital of Wisconsin, Milwaukee, Wisconsin

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

ABSTRACT

OBJECTIVE. Our objective was to determine whether parent-reported, high-quality primary care was associated with decreased nonurgent pediatric emergency department utilization.

METHODS. A retrospective analysis of prospectively collected data for a cohort of children from the 2000–2001 and 2001–2002 Medical Expenditure Panel Survey panels was performed. Baseline parent-reported quality of primary care with respect to family-centeredness, timeliness, and realized access (a measure of the child’s ability to receive necessary care and referrals) was assessed by using composite scores from the Consumer Assessment of Healthcare Providers and Systems survey. The primary outcomes were the numbers of subsequent nonurgent and urgent emergency department visits per child.

RESULTS. Of 8823 children included, 70.0% rated family-centeredness, 88.2% rated realized access, and 55.6% rated timeliness as high quality. After adjustment for demographic factors and health status, high-quality family-centeredness was associated with a 42% reduction in nonurgent emergency department visits for publicly insured children and a 49% reduction for children ≤2 years of age. Greater realized access was associated with a 44% reduction in nonurgent emergency department visits for children 3 to 11 years of age and a 56% reduction for children ≥12 years of age. Greater realized access was also associated with decreased nonurgent emergency department visits for publicly and privately insured children (37% and 35%, respectively). There was no significant association between timeliness and nonurgent emergency department utilization, nor was any quality-of-care domain associated with urgent emergency department utilization.

CONCLUSIONS. Parent-reported, high-quality family-centeredness and a high level of realized access to primary care were associated with decreased subsequent nonurgent emergency department visits for children. Parent reports of health care quality in these domains provide important complementary information on health care quality.
Evidence shows that there is an association between quality of primary care and ED utilization, with lower quality of care being associated with increased utilization. However, the quality of care evaluated to date focuses on processes of care, such as the timely receipt of immunizations or the prescribing of asthma controller medications. Although these process-of-care measures are important, we think that they fall short of a comprehensive assessment of health care quality, and we postulate that patient-reported quality domains might provide complementary information for assessment of health care quality.

Of the 6 quality-of-care domains (safety, effectiveness, patient- or family-centeredness, timeliness, efficiency, and equity) highlighted for improvement in the Institute of Medicine report *Crossing the Quality Chasm*, many are domains for which patient assessment is an important measure. The authors of the Institute of Medicine report thought that “patients’ experiences should be the fundamental source of the definition of quality”; however, patient experiences with care have largely been unexplored, in part because of the lack of means for measuring patient-reported quality of care for these domains.

The development of the Consumer Assessment of Healthcare Providers and Systems (CAHPS) (formerly the Consumer Assessment of Health Plans Study), with a parent-administered survey to evaluate quality of care for children, provides an instrument that assesses patient experiences with care validly. The CAHPS survey specifically addresses, from the patient or parent perspective, the family-centeredness, timeliness, and certain aspects of the effectiveness of care received from a primary care provider.

The objective of the study was to determine the association between parent-reported quality of care in specific domains and subsequent ED utilization by children. We hypothesized that high-quality primary care, as indicated by parent-reported family-centeredness, timeliness, and effectiveness, would be associated with a decrease in subsequent nonurgent pediatric ED visits, without an associated decrease in urgent ED visits.

**METHODS**

**Study Design and Source of Data**
This was a retrospective analysis of prospectively collected data for the cohort of children (<17 years of age) in the 2000 to 2001 (panel 5) and 2001 to 2002 (panel 6) Medical Expenditure Panel Survey (MEPS) panels. MEPS, a subsample of the National Health Interview Survey, is a nationally representative survey of the US civilian noninstitutionalized population conducted annually by the Agency for Healthcare Research and Quality. Data are collected on the specific health services that US residents use and how frequently they use them. MEPS uses a longitudinal system of 5 computer-assisted, in-person interviews over a 2-year period to collect household data, including demographic features, perceived health status, screener questions for children with special health care needs, insurance status, and ED, primary care, and hospital utilization. In 2000, MEPS added CAHPS questions to the second round of interview questions, which allowed analysis of parent-reported quality of care and subsequent child health care utilization in the same national database.

**Study Population**
All children for whom a caregiver completed the CAHPS questions (>96% response rate) were eligible. Children were excluded if complete follow-up information was not available for the subsequent 3 rounds of interviews; this would occur if the child left the household, was institutionalized, died, or joined the military.

**Outcome Measures**
The primary outcomes were the numbers of parent-reported nonurgent and urgent ED visits per child. Visit urgency was determined on the basis of the parent-reported resources used during the visit. This method of assigning urgency was used previously in large database studies and is the database research method preferred by the Ambulatory Pediatric Association Pediatric Emergency Medicine Special Interest Group. Exactly as in a previous study, any visit that resulted in laboratory work, a radiograph, an electrocardiogram, an electroencephalogram, or admission to the hospital was considered urgent; all others were classified as nonurgent. The dates for all ED visits were recorded, and only those that occurred after the date of completion of the second interview (when the CAHPS responses were collected) were included in the analysis; this ensured that all ED visits occurred after completion of the CAHPS survey.

**Assessment of Health Care Quality Domains**
The quality of primary care received by the child was assessed with the CAHPS questions. The 9 quality-of-care questions were grouped into composite scores for the 3 quality domains, that is, (1) family-centeredness,
(2) timeliness, and (3) “realized access” to care.\textsuperscript{35} Realized access is an aspect of health care effectiveness that addresses specifically whether families perceive problems in obtaining necessary care or referrals. The individual questions constituting each composite, and their Likert scales, are listed in Table 1.

The composite scores for each child were obtained by averaging the individual answered questions that constitute the quality-of-care composites, with higher scores corresponding to higher quality. These composite scores and question groupings were developed by the Agency for Healthcare Research and Quality, with 9 of the 10 original individual questions being added to the MEPS survey. To be consistent with previous research, the timeliness and family-centeredness composites, both scored on a 4-point Likert scale, were dichotomized, with scores of >3.5 indicating the highest-quality care and scores of ≤3.5 indicating lower-quality care.\textsuperscript{36} For realized access, scored on a 3-point Likert scale, scores were dichotomized as 3 vs <3.

Other patient/family characteristics extracted from the MEPS data set included age (categorized as 0–2 years, 3–11 years, or 12–17 years), gender, race/ethnicity (non-Latino white, Latino, non-Latino black, or other, which included Asian/Pacific Islander and Native American/Alaskan native), having a usual source of care (yes or no), having special health care needs (yes or no), interview language (English, Spanish, or both), and parent-reported child health status (excellent or very good versus good, fair, or poor). Multiple indicators of socioeconomic status were included, that is, combined annual family income categorized as a percentage of the federal poverty level and dichotomized as poor or nearly poor (<125% of the federal poverty level) versus higher than nearly poor (≥125% of the federal poverty level), health insurance (public, private, or none), and the highest level of parental educational attainment (dichotomized as less than high school graduate versus high school graduate or beyond).\textsuperscript{37}

### TABLE 1: Individual CAHPS Questions Constituting the Quality-of-Care Composites

<table>
<thead>
<tr>
<th>Family-centeredness (1 = never, 2 = sometimes, 3 = often, 4 = always)</th>
</tr>
</thead>
<tbody>
<tr>
<td>How often providers explained things so you understood</td>
</tr>
<tr>
<td>How often providers showed respect for what you had to say</td>
</tr>
<tr>
<td>How often providers spent enough time with you</td>
</tr>
<tr>
<td>How often providers listened carefully to you</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Timeliness (1 = never, 2 = sometimes, 3 = often, 4 = always)</th>
</tr>
</thead>
<tbody>
<tr>
<td>How often you got an appointment for an injury or illness as soon as you wanted</td>
</tr>
<tr>
<td>How often you got a routine appointment as soon as you wanted</td>
</tr>
<tr>
<td>How often you were able to get help by telephone</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Realized access (1 = a big problem, 2 = somewhat of a problem, 3 = not a problem)</th>
</tr>
</thead>
<tbody>
<tr>
<td>How big a problem it was to get care you or a doctor thought was necessary</td>
</tr>
<tr>
<td>How big a problem it was to get a referral to a specialist</td>
</tr>
</tbody>
</table>

Highest quality was indicated by a composite score of >3.5 on the 4-point scale for family-centeredness and a score of 3 on the 3-point scale for realized access.

### Analyses

Two separate longitudinal cohorts of children were combined for the study, to provide sufficient power to allow for age and insurance subgroup analyses. The 2 cohorts included were the first 2 cohorts within MEPS to have CAHPS responses. Analyses were weighted to yield national estimates for civilian noninstitutionalized children in the United States. The “survey” commands of Stata 9.1 (Stata Corp, College Station, TX) were used for weighting the means and proportions and for the Poisson regression analyses, accounting for the nonrandom sampling strategy used in MEPS.\textsuperscript{31}

Comparisons between the percentage of the population and the percentage of utilization by each demographic subgroup were made by using multiple-comparison-adjusted, survey-weighted tests. Generalized Poisson regression analyses, allowing for incorporation of an additional variance parameter to account for potential overdispersion, were used to assess the association between each composite and the number of subsequent nonurgent and urgent ED visits per child. Results from the regression analyses are presented as the incidence rate ratios (IRRs) for the number of urgent or nonurgent ED visits per child. An IRR of <1 indicates that higher quality is associated with fewer ED visits per child.

Each quality-of-care domain was analyzed in a multivariate, generalized, Poisson regression analysis with each measure of ED utilization. All covariates were entered into the model, grouped as described previously and as listed in Table 2. The groupings were derived by collapsing categories for a covariate when the relationships with ED utilization were similar. Because there were significant interactions between (1) insurance type and the composites and (2) age and the composites, stratified analyses were completed for insurance type (private insurance or public insurance) and age. The stratified regression results for children with no insurance are not presented because of small sample size.

### RESULTS

There were 8823 children eligible for the study. Analysis of ED visits revealed that 1419 children (16.1%) made a total of 1786 ED visits during the study period; 768 (43%) of those visits were classified as nonurgent. The numbers of nonurgent visits ranged from 1 nonurgent visit for 500 children to 10 nonurgent visits for 1 child.

The characteristics of the children and the percentages of children with ≥1 urgent ED visit and ≥1 nonurgent ED visit according to characteristic are shown in Table 2. The youngest children, for example, while constituting only 17.3% of the population, accounted for 17.9% of children making urgent ED visits but 34.8% of children making nonurgent ED visits. Therefore, the youngest children were overrepresented with respect to children making nonurgent ED visits, whereas older
children were underrepresented. Children with public insurance, while constituting only 25.0% of the population, were overrepresented with respect to both urgent and nonurgent ED use, and uninsured children were underrepresented with respect to urgent ED use. Other characteristics associated with a disproportionately large number of children with nonurgent ED visits included being male, being poor/near poor, having special health care needs, having worse parent-reported health status, having a usual source of care, and having a lower degree of family educational attainment.

Assessment of parent-reported, quality-of-care domains showed that 70% of parents rated the family-centeredness of their child’s primary care as high quality, approximately one half rated timeliness as high quality, and almost 90% rated realized access as high quality (Table 3). For both family-centered and realized access, high-quality care was associated with a lower num-

<table>
<thead>
<tr>
<th>TABLE 2 Characteristics of the Study Population and Associated Proportions of Children With Urgent and Nonurgent ED Visits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
</tr>
<tr>
<td>Age, y</td>
</tr>
<tr>
<td>0–2</td>
</tr>
<tr>
<td>3–11</td>
</tr>
<tr>
<td>12–17</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Race/ethnicity</td>
</tr>
<tr>
<td>Non-Latino white</td>
</tr>
<tr>
<td>Non-Latino black</td>
</tr>
<tr>
<td>Latino</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Insurance</td>
</tr>
<tr>
<td>Private</td>
</tr>
<tr>
<td>Public</td>
</tr>
<tr>
<td>Uninsured</td>
</tr>
<tr>
<td>Income</td>
</tr>
<tr>
<td>&lt;200% of poverty level</td>
</tr>
<tr>
<td>≥200% of poverty level</td>
</tr>
<tr>
<td>Child with special health care needs</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>Reported health status of child</td>
</tr>
<tr>
<td>Excellent/very good</td>
</tr>
<tr>
<td>Good/fair/poor</td>
</tr>
<tr>
<td>Child identified as having a usual source of care</td>
</tr>
<tr>
<td>Yes</td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>Highest family education</td>
</tr>
<tr>
<td>High school graduate or less</td>
</tr>
<tr>
<td>More than high school</td>
</tr>
<tr>
<td>Interview language</td>
</tr>
<tr>
<td>English</td>
</tr>
<tr>
<td>Spanish</td>
</tr>
<tr>
<td>English and Spanish</td>
</tr>
<tr>
<td>Other</td>
</tr>
</tbody>
</table>

Percentages were derived from weighting to yield national estimates from the sample of 8823 children.

* Significantly different from percentage of population, P < .01.

<table>
<thead>
<tr>
<th>TABLE 3 Weighted Percentages of Parents Reporting the Highest Quality of Care for Their Children With Associated Percentages of Children With Urgent and Nonurgent ED Visits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality-of-Care Composite, Highest Quality</td>
</tr>
<tr>
<td>Family-centeredness</td>
</tr>
<tr>
<td>Timeliness</td>
</tr>
<tr>
<td>Realized access</td>
</tr>
</tbody>
</table>

Highest quality indicates a composite score of >3.5 on the 4-point scale for family-centeredness and timeliness and a score of 3 on the 3-point scale for realized access.

* Significantly different from percentage of population, P < .01.
number of children with ≥1 nonurgent ED visit. For timeliness, the percentages of children with urgent and nonurgent visits were similar, with the percentage of children with urgent visits being significantly less than the percentage reporting high-quality care.

In univariate regression analyses, high-quality family-centered care (IRR: 0.64; 95% confidence interval [CI]: 0.50–0.82) and greater reports of realized access (IRR: 0.53; 95% CI: 0.37–0.78) were both associated with significant decreases in subsequent nonurgent ED visits per child. An IRR of 0.64 translates into a 36% reduction in nonurgent ED visits, with an IRR of 0.53 being equal to a 47% reduction in visits. There was no association between timeliness and nonurgent ED visits, nor were any of the 3 quality-of-care domains associated with decreased numbers of subsequent urgent ED visits (not shown).

To adjust for potential confounding by child and family characteristics, a multivariate, generalized, Poisson regression model was constructed. Because of significant interactions between the quality-of-care domains and both insurance type and age, the analysis was stratified according to those factors. Analysis of publicly insured children revealed that those with high-quality family-centeredness and realized access had 42% and 37% decreases in nonurgent ED visits per child, respectively (shown in Table 4 by IRRs with CIs not including 1). In contrast, high-quality timeliness was not associated with children making urgent or nonurgent visits. Among children with private insurance (Table 4), only high-quality realized access was associated with fewer nonurgent ED visits by children. High-quality care was not associated with decreased urgent ED utilization in any domain.

Analyses stratified according to age (Table 5) again showed no significant association between the specific quality-of-care domains and urgent ED visits. Among the youngest children, high-quality family-centered care was associated with a substantial decrease in nonurgent ED visits per child; for older children, greater realized access was associated with fewer nonurgent ED visits (Table 5).

In addition to the quality-of-care composites, 2 other factors, namely, being a child with special health care needs and worse parent-reported child health status, were associated consistently with ED utilization in all multivariate models. In the model for children with public insurance coverage, being a child with special health care needs was associated with increases in both urgent (IRR: 2.07; 95% CI: 1.51–2.86) and nonurgent (IRR: 1.73; 95% CI: 1.39–2.64) ED visits. Worse parent-reported child health status was associated only with an increase in nonurgent ED visits (IRR: 2.27; 95% CI: 1.32–3.93). These associations were similar in other models. None of the other characteristics, including gender, race/ethnicity, income, having a usual source of care, parental educational attainment, and interview language, showed consistent association with ED utilization.

**DISCUSSION**

The study findings indicated that parent-reported, high-quality, family-centered care and increased realized access were both associated with decreased numbers of subsequent nonurgent ED visits. For the 2 groups known from previous research to have disproportionately high ED utilization, that is, children ≤2 years of age and those with public insurance, high-quality family-centered care was associated with 40% to 50% reductions in subsequent nonurgent ED utilization, as evidenced by IRRs of 0.51 and 0.58, respectively. For older children and for children with private insurance, realized access (the ability to obtain referrals or other care that the doctor or family thought was necessary) was associated most strongly with decreased nonurgent ED visits. There were no statistically significant associations between any of the quality-of-care composites and urgent ED utilization, as evidenced by the CIs including 1.0. This lack of an association was expected, given that care for urgent conditions is less discretionary.

The association between high-quality family-centered care and nonurgent ED utilization for the youngest children and for publicly insured children is of interest.

### Table 4: Associations Between Quality-of-Care Composites and Urgent and Nonurgent ED Visits, Stratified According to Insurance Type

<table>
<thead>
<tr>
<th>Quality-of-Care Composite</th>
<th>Urgent ED Visits (IRR 95% CI)</th>
<th>Nonurgent ED Visits (IRR 95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Public insurance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family-centeredness</td>
<td>0.83 (0.61–1.13)</td>
<td>0.58 (0.40–0.85)</td>
</tr>
<tr>
<td>Timeliness</td>
<td>0.95 (0.69–1.29)</td>
<td>0.74 (0.53–1.02)</td>
</tr>
<tr>
<td>Realized access</td>
<td>0.97 (0.70–1.34)</td>
<td>0.63 (0.42–0.94)</td>
</tr>
<tr>
<td><strong>Private insurance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family-centeredness</td>
<td>1.08 (0.85–1.39)</td>
<td>0.81 (0.59–1.12)</td>
</tr>
<tr>
<td>Timeliness</td>
<td>0.92 (0.75–1.13)</td>
<td>0.87 (0.67–1.12)</td>
</tr>
<tr>
<td>Realized access</td>
<td>0.96 (0.68–1.34)</td>
<td>0.65 (0.43–0.99)</td>
</tr>
</tbody>
</table>

A multivariate, generalized Poisson regression analysis adjusted for age, gender, race/ethnicity, income, special health care needs, reported health status, existence of a usual source of care, family education, and interview language was performed. Children with no insurance are not presented because of the small sample size.
Children ≤2 years of age develop more acute illnesses and therefore access the health care system more frequently, for both primary care and ED visits, than do older children. When nonurgent conditions arise among these youngest children, parents who think that their children have access to high-quality family-centered care (primary care providers who respect what the parents have to say, who listen carefully, and who spend enough time with the children) may choose to have the children seen by that provider, rather than by a stranger in the ED. The content of the high-quality communication between the parents and the primary care providers may also play a role, with providers who practice family-centered care possibly explaining to families when ED care is warranted and when it is better to wait to be seen.

For children with public insurance, the issue of access to other sources of care becomes paramount. When transportation issues or financial difficulties make accessing the primary care provider more difficult, parents may use the ED for nonurgent conditions unless they perceive the loss of continuity with the primary care provider to be detrimental. The bond with the provider who delivers high-quality family-centered care may be strong enough to offset the short-term advantage of convenient ED care.

Nonurgent ED visits were less common for older children and children with private insurance than for younger children and children with public insurance. It seems that parental perceptions that the provider is not able to meet the needs of the child are more important in determining to seek ED care than the bond between the provider and the family. Whether this results from a provider actually not being able to meet the needs of the child or is simply the parent’s perception of the provider’s inability cannot be determined from this study. In either case, realized access was associated strongly with nonurgent utilization in these populations.

We had hypothesized that timely access to primary care would be associated with decreased nonurgent ED use, but we found no significant associations. Two explanations may be operative. First, the timeliness composite is composed of questions that assess care across many domains, including well-child care, sick care, and after-hours care, only some of which may be related to nonurgent ED utilization. Second, it may be that timely access alone is not sufficient to decrease ED utilization but the past interactions and content of previous discussions between the provider and the family dominate the family’s decision to seek ED care.

The lack of an association between parent-reported quality of care and urgent ED visits also warrants exploration. As hypothesized, the quality of interactions with the primary care provider had no significant association with urgent ED visits. Studies that rely on a decrease in total ED utilization (approximately one half of which may not be discretionary) as the sole outcome may miss important reductions in nonurgent visits and may assume incorrectly that proposed interventions are not effective.

In addition to the quality domains evaluated, other covariates were associated with increased nonurgent ED visits. The findings that worse reported health status, having special health care needs, and having lower income were associated with increased ED visits are consistent with previous studies. The bivariate finding that having a usual source of care was associated with increased nonurgent ED visits is not as intuitively clear. One possible explanation is that children who lack insurance frequently lack a usual source of care; this population utilizes health care, including the ED, less frequently than do children with insurance. This explanation is supported by the finding that only items relating to an increased level of baseline illness (health

### Table 5

**Associations Between Quality-of-Care Composites and Urgent and Nonurgent ED Visits, Stratified According to Age**

<table>
<thead>
<tr>
<th>Quality-of-Care Composite</th>
<th>0–2 y IRR (95% CI)</th>
<th>3–11 y IRR (95% CI)</th>
<th>≥12 y IRR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family-centeredness</td>
<td>0.70 (0.48–1.01)</td>
<td>1.01 (0.76–1.35)</td>
<td>1.09 (0.78–1.52)</td>
</tr>
<tr>
<td>Timeliness</td>
<td>0.71 (0.50–1.02)</td>
<td>0.93 (0.72–1.20)</td>
<td>1.02 (0.78–1.34)</td>
</tr>
<tr>
<td>Realized access</td>
<td>0.87 (0.55–1.38)</td>
<td>1.07 (0.78–1.46)</td>
<td>0.92 (0.62–1.35)</td>
</tr>
</tbody>
</table>

A multivariate, generalized Poisson regression analysis adjusted for insurance type, gender, race/ethnicity, income, special health care needs, reported health status, existence of a usual source of care, family education, and interview language was performed.
status and special health care needs) remained significant after adjustment.

Certain study limitations should be noted. First, as with all MEPS data, the family reported the number of ED visits; there was no independent verification, which allowed for possible inaccurate reporting of data. The MEPS data, however, are collected prospectively, and the results are population based, with a large sample size. The 96% response rate for the CAHPS survey also supports the generalizability of the results. Second, the definition of urgency of ED visits was based on the resources used during the ED visit. Therefore, some visits might have been misclassified, especially those for conditions such as asthma exacerbation, for which therapeutic rather than diagnostic interventions would have resulted in classification as urgent. Classification based on resource utilization has been used in previous research, with criteria identical to those used in this study. Finally, we did not have a sufficient sample size for uninsured children to address adequately the association between quality of care and ED utilization in this population.

In the case of pediatric ED utilization, it is the parent who decides when and where to seek care for a child. Therefore, it is only by understanding the parent's perspective on the relationship between the family and the primary care provider that one can understand the decisions that families make when deciding where and when to seek care. The assessment of patient- or parent-reported quality of care, although not itself a comprehensive measure of health care quality, provides important complementary information about the quality of care that children receive. In addition, parent-reported quality of care in these specific domains may aid in the identification of children at increased risk for nonurgent ED visits and may allow for targeted interventions aimed at improving the care that such children receive.

CONCLUSIONS
This is the first study, to our knowledge, to show that parent-reported, high-quality family-centeredness and ability to receive needed care are associated with decreased nonurgent ED visits by children. We think that the addition of these parent-reported measures of child health care quality to the process or outcome measures of quality yields a more-comprehensive measure of the quality of care that children receive, at least with regard to risk factors for nonurgent ED utilization, and can aid in the identification of children and families receiving lower-quality care.

ACKNOWLEDGMENT
This work was supported in part by grant K08 HS015482-01A1 (to Dr Brousseau) from the Agency for Healthcare Research and Quality.

REFERENCES
20. Adams RJ, Fuhlbrigge A, Finkelstein JA, et al. Impact of inhaled antiinflammatory therapy on hospitalization and emer-


33. Mistry RD, Hoffmann RG, Yauck JS, Brousseau DC. Association between parental and childhood emergency department utilization. *Pediatrics*. 2005;115(2). Available at: www.pediatrics.org/cgi/content/full/115/2/e147


The Host Hospital 24-Hour Underreferral Rate: An Automated Measure of Call-Center Safety

Daniel A. Hirsh, MDa,b,c, Harold K. Simon, MD, MBAa,b,c, Robert Masseyd, Lisa Thornton, RN,e, Joseph E. Simon, MS, MDc,e

Departments of aPediatrics and bEmergency Medicine, Emory University, Atlanta, Georgia; Departments of cPediatrics, dInformation Systems and Technology, and eEmergency Services, Children’s Healthcare of Atlanta, Atlanta, Georgia

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

ABSTRACT

OBJECTIVES. The goals were to (1) define and illustrate an automated method of monitoring the safety of telephone triage, (2) demonstrate that this method approximates reasonably a more-global safety measure, and (3) describe the month-to-month variability of this automated measure for the call center studied.

METHODS. From October 2005 through March 2006, hospitalizations at a tertiary care pediatric hospital after calls to its call center were matched with their respective call-center dispositions. The host hospital 24-hour underreferral rate was defined as the percentage of total admissions to the study institution within 24 hours after a call to the call center for treatment of the same illness or injury that had been assigned a nonurgent disposition by the call center. A convenience sample of call-center calls was surveyed for admissions to other facilities. This sample was then combined with admissions to the pediatric hospital to estimate a true 24-hour underreferral rate. Underreferrals were subjected to clinical and statistical analyses.

RESULTS. The host hospital 24-hour underreferral rate was 5.2%. The estimated true 24-hour underreferral rate was 5.95% ± 2.75%. Diagnoses frequently associated with underreferral were gastroenteritis, croup, asthma, and bronchiolitis. Underreferred patients admitted to the study institution were hospitalized for an average of 1.6 ± 1.1 days, compared with 2.8 ± 3.1 days for patients referred by the call center to a higher level of care. The monthly SD of the host hospital 24-hour underreferral rate was 1.56%.

CONCLUSIONS. For the call center studied, the host hospital 24-hour underreferral rate could be determined easily and objectively and approximated reasonably the true 24-hour underreferral rate. The month-to-month variability of the host hospital 24-hour underreferral rate was sufficiently small to allow for meaningful internal trending analyses.
Pediatric telephone triage centers carry a heavy burden of responsibility. Failing to refer a sick child to a higher level of care (underreferral) could severely compromise patient safety. A benchmark measure of the tendency of a call center (CC) to underrefer would be highly desirable. This benchmark should be determined easily, allowing a CC to use serial measurements of this benchmark to identify its underreferral trend as changes are made to protocols, policies, and training procedures. For trending of this benchmark to be meaningful, its variability from month to month should be low and the calls responsible for any trend should be identified easily.

Three studies have examined a CC’s tendency to underrefer pediatric patients.1–3 Lee et al. prospectively compared calls managed by a CC and calls managed by pediatricians. Kempe et al. used postcall telephone surveys to determine underreferral rates. Those 2 studies are worthy contributions to the telephone triage literature. However, their methods do not lend themselves easily to serial measurements of CCs’ underreferral rates.

A second study by Kempe et al. linked CC data with the claims information for a closed, managed-care population. This data-mining approach allowed Kempe et al. to study 32,968 calls, with 378 patients admitted within 24 hours after calling the CC. Fifty-six of the 378 admissions had been provided with a nonurgent disposition by the CC. Kempe et al. chose to define the potential underreferral rate for the CC studied by using total calls as the denominator. Specifically, the potential underreferral rate was defined by Kempe et al. as being 56 divided by 32,968, or 0.17% (1 of every 589 calls).

Use of total calls as the denominator in defining a CC’s underreferral rate necessitates that total calls be the basis of study. Without electronic access to insurance claims data for at least a subpopulation of callers, this forces a CC to rely on a follow-up call method, with its inherent costs and volume limitations, when studying its tendency to underrefer. The current study was undertaken to demonstrate the use of an automated measure of a CC’s tendency to underrefer, based on patients admitted within 24 hours to a single hospital, and to demonstrate that it approximates reasonably a more-global measure of that CC’s tendency to underrefer.

METHODS

Study Site

From October 2005 through March 2006, records from a tertiary care pediatric hospital system (Children’s Healthcare of Atlanta, Atlanta, GA) and its associated CC were reviewed. Children’s Healthcare of Atlanta inpatient facilities consist of 2 geographically separated, pediatric hospitals. These 2 hospitals account for 68% of all pediatric admissions in the 7-county metropolitan service area. The CC manages calls from both pediatrician answering services (>600 enrolled physicians) and a community telephone line (250-KIDS). Eighty-five percent of all patients calling the CC are patients of enrolled pediatricians. Staff membership at Children’s Healthcare of Atlanta is not a requirement for a pediatrician to access CC services. The service is free to enrolled pediatricians. The CC uses 70 computerized, branched-chain, logic protocols administered by registered nurses. Each nurse undergoes a 2-month orientation, including didactic sessions, call observation, and call management in the presence of an experienced CC nurse, before being allowed to manage calls independently.

Daily Query of the CC and Inpatient Databases

Although the CC and inpatient databases at Children’s Healthcare of Atlanta are not interfaced, the Children’s Healthcare of Atlanta Department of Information Systems and Technology was able to create a daily query of these databases that identified, for each call to the CC, matches (according to birth date and last name) that appeared within 24 hours after the call as Children’s Healthcare of Atlanta inpatients. During the study period, this report was prepared each day for the calls placed 4 days earlier. For example, on the tenth of the month, a report of all admissions within 24 hours after calls placed on the sixth of the month was available for analysis. Additional information listed in the reports for each call and admission included call and admission dates and times, CC nurse, CC disposition, referral justification (if referred), CC protocol used, admitting diagnosis, CC identification number, and inpatient medical charts number.

Urgent Referral Rate

A protocol branch might end in one of many different dispositions. These dispositions were divided into 3 groups: “911/ED/AHC,” “call another provider,” and “nonurgent.” The 911/ED/AHC group represented all 911 referrals and all referrals to an emergency department (ED) or another after-hours care (AHC) setting for examination. The call another provider group included all dispositions that directed the caller to speak promptly to another provider, where promptly was defined as <4 hours. “Another provider” included pediatricians, subspecialty physicians, and the Georgia Regional Poison Control Center. The nonurgent group included all remaining dispositions not involving prompt examination or prompt telephone contact with another provider. For example, it included the commonly used disposition “call pediatrician in morning.” The urgent referral rate for a given time period was defined as the number of 911/ED/AHC dispositions expressed as a percentage of the total number of advice calls managed during that time period.

Many pediatricians who subscribe to CC services request to perform secondary triage for any patient judged to warrant an ED referral by the CC nurse. The disposi-
tion of record for a given call, for purposes of calculating the CC urgent referral rate, was the disposition determined to be most appropriate by the CC nurse. For example, if a call was sent to a physician because of his or her standing request to perform secondary triage for ED referrals, then this call was regarded as an ED referral.

Underreferral Rate Calculations
The host hospital 24-hour underreferral rate was defined as the number of children provided with a nonurgent disposition who were subsequently admitted to Children’s Healthcare of Atlanta within 24 hours for treatment of the same illness or injury, expressed as a percentage of total calls resulting in admission to Children’s Healthcare of Atlanta within 24 hours for treatment of the same illness or injury. This rate, based on hospitalizations at Children’s Healthcare of Atlanta, serves as a proxy for the more-global rate including hospitalizations at any institution. In particular, the “true” 24-hour underreferral rate would be the number of children provided with a nonurgent disposition who were subsequently admitted to any hospital within 24 hours for treatment of the same illness or injury, expressed as a percentage of total calls resulting in admission to any hospital within 24 hours for treatment of the same illness or injury. To determine the degree to which the host hospital 24-hour underreferral rate approximates the true 24-hour underreferral rate, a convenience sample of calls to the CC underwent follow-up monitoring via telephone during the study period. This convenience sample was designed to avoid disproportionate sampling of a particular CC shift or day of the week. The rate of admission to a hospital other than Children’s Healthcare of Atlanta after a call to the CC, for treatment of the same illness or injury, was then estimated by using standard sample size calculations, with a confidence interval (CI) of 95%. In a similar manner, the sampling results were then used to estimate the number of children admitted to another hospital within 24 hours after a nonurgent CC disposition. These results were then combined with the host hospital 24-hour underreferral rate data.

Calls with the disposition “information only” were eliminated from consideration in analyses of underreferral rates. This disposition is used when a CC nurse does not feel comfortable providing definitive triage advice. For example, this disposition is used routinely when the caller is not with the child.

Analysis of Underreferrals
The frequency with which a given protocol resulted in underreferral was determined and compared with the overall use rate for that protocol during the study. Final diagnoses, lengths of stay, age distributions, and other characteristics of underreferred patients were determined.

Statistical Methods and Institutional Review Board Approval
Categorical variables were analyzed by using \( \chi^2 \) analysis. Sampling errors were estimated by using standard sample size calculations. This study was approved by the institutional review boards of both Children’s Healthcare of Atlanta and Emory University.

RESULTS

Disposition Rates
The total number of patient contacts for the study period was 139,621. Calls eliminated because a definitive disposition was not assigned included 6535 callers who declined the need for advice when contacted, 86 calls for which a communication barrier could not be overcome, and 6028 callers who were provided with information only, according to protocol or CC policy. This left 126,972 callers during the study period who were provided with definitive dispositions. The urgent referral rate was 24.3%. The call another provider rate was 20.7%. The remaining 55% of callers were provided with a nonurgent disposition.

Twenty-Four–Hour Underreferral Rate
During the study period, 831 children were admitted to Children’s Healthcare of Atlanta within 24 hours after contact with the CC, of whom 807 were judged to have been admitted for treatment of the same illness or injury that prompted the call to the CC. Forty-six of the 831 children had received a nonurgent disposition from the CC. Forty-two of those children were judged to have been admitted for treatment of the same illness or injury that prompted the call to the CC, for a host hospital 24-hour underreferral rate of 5.2% (42 of 807 children). With elimination of the subjective determination of the relationship of illness prompting the call to illness leading to admission, the 24-hour Children’s Healthcare of Atlanta underreferral rate was 5.5% (46 of 831 children).

During the same study period, a convenience sample of 4986 callers to the CC was studied. Thirteen admissions to a non–Children’s Healthcare of Atlanta inpatient facility within 24 hours after contact with the CC were identified. (One call was eliminated from consideration because admission time could not be determined reliably; that caller had been provided with a nonurgent disposition.) Eight of the 13 calls had been provided with an urgent referral disposition by the CC; 5 had been directed to call another provider. A simple extrapolation of these 13 admissions from the sample of 4986 to 126,972 total calls during the study period resulted in an
estimate of 331 admissions to a non–Children’s Healthcare of Atlanta inpatient facility \[126972 \times (13/4986)\]. By treating these 4986 calls as a sample of the 126972 calls and using a 95% CI, the possible range of the number of admissions to other hospitals after a call to the CC during the study period was determined to be 142 to 520 admissions (Table 1).

None of the 13 admissions to other hospitals received a nonurgent disposition from the CC. In Table 1, in a manner identical to that described above, the 95% CI for this result was determined when applied to the low, mean, and high estimates of admissions to other facilities. Finally, in Table 1, these data were combined with admissions to Children’s Healthcare of Atlanta to calculate a true 24-hour underreferral rate of 5.95 ± 2.75% (95% CI).

With the use of total calls as the denominator in the formula defining the CC underreferral rate (the method used in the studies by Kempe et al\(^2,3\)), this rate would range from 1 of 1094 calls to 1 of 3023 calls (95% CI). Stated as a percentage, this rate range is 0.03% to 0.09% (95% CI). The monthly Children’s Healthcare of Atlanta underreferral rate during the 6 months studied varied from 3.5% to 7.9% (mean: 5.39%; SD: 1.56%).

**Clinical Characteristics of Underreferrals**

Table 2 lists the final diagnoses for the 42 underreferrals to Children’s Healthcare of Atlanta. Gastroenteritis, croup, asthma, and bronchiolitis accounted for 66% of the cases. The average length of stay was 1.6 days (SD: 1.1 days), compared with 2.8 days (SD: 3.1 days) for patients referred to a higher level of care (\(P = .016\)). Only 1 underreferred patient was hospitalized for >3 days (6-day admission for management of an infected branchial cleft cyst). Two underreferred patients were admitted to the ICU for a total of 2 ICU days, one with asthma and the other with new-onset diabetes mellitus. The age distribution for underreferrals was not significantly different from the overall age distribution for cases managed by the CC, as determined with an F ratio (analysis of variance) of 1.6 in an age group analysis with 4 age groups.

**Protocols Used to Manage Underreferrals**

Comparisons of the frequency of protocol use for underreferrals and the overall frequency of use of a given protocol at the CC yielded no statistically significant differences, except for the croup protocol (odds ratio: 9.83; \(P < .001\)).

| TABLE 1 Calculation of the Range of 24-Hour Underreferral Rates by Combining Data From Children’s Healthcare of Atlanta and the Sample of Admissions to Non–Children’s Healthcare of Atlanta Facilities |
|-----------------|-----------------|-----------------|----------------- |
| Row | Parameter | Low Estimate | Extrapolated Number | High Estimate |
| 1 | Non–Children’s Healthcare of Atlanta admissions within 24 h, \(n\) | 142 | 331 | 520 |
| 2 | Children’s Healthcare of Atlanta admissions within 24 h, \(n\) | 807 | 807 | 807 |
| 3 | Total admissions, sum of rows 1 and 2, \(n\) | 949 | 1138 | 1327 |
| | Low | Low Estimate | Extrapolated Number | High Estimate |
| 4 | Non–Children’s Healthcare of Atlanta admissions within 24 h after nonurgent disposition, \(n\) | 0 | 0 | 19 |
| 5 | Children’s Healthcare of Atlanta admissions within 24 h after nonurgent disposition, \(n\) | 42 | 42 | 42 |
| 6 | Sum of rows 4 and 5, \(n\) | 42 | 42 | 61 |
| 7 | Ratio of rows 6 and 3, \(n\) | 4.4 | 4.4 | 6.4 |

\(a\) Values were derived by using a 0.3% expected response ratio (rate of admission to a non–Children’s Healthcare of Atlanta facility) and a 95% CI. Specifically, 0.3% was calculated by using the ratio \(x/0.65 = 32/68\), where 68% is the Children’s Healthcare of Atlanta market share, 32% is the non–Children’s Healthcare of Atlanta market share, and 0.65% is the rate of admission to Children’s Healthcare of Atlanta within 24 hours after a call to the CC.

\(b\) Values were derived by using a 7.5% expected response ratio of surveyed/admitted patients who were underreferred and a 95% CI.

| TABLE 2 Diagnoses of Admissions to Children’s Healthcare of Atlanta Within 24 Hours After a Nonurgent Disposition by the CC |
|-----------------|-----------------|
| Diagnosis | Frequency, \(n\) (%) |
| Acute gastroenteritis | 9 (21.4) |
| Croup | 8 (19) |
| Asthma | 6 (14.3) |
| Bronchiolitis | 5 (11.9) |
| Newborn fever (neither infant had fever at time of call) | 2 (4.7) |
| Cellulitis/abscess | 2 (4.7) |
| Vomiting, pneumonia, idiopathic thrombocytopenic purpura, periorbital cellulitis, gastric tube displacement, new-onset diabetes mellitus, esophageal foreign body, appendicitis (nonperforated), staphylococcal scalded skin syndrome, and upper respiratory infection | 1 (2) each |
DISCUSSION

Underreferral and overreferral are both issues in pediatric telephone triage. Several studies support the view that pediatric CCs overrefer callers for urgent examination, at least when the referrals are evaluated retrospectively by examining physicians or expert panels.4–7 The question of whether a CC can safely reduce its urgent referral rate without compromising its underreferral rate arises from these studies. Designing studies to answer this question requires an easily obtainable and objective measure of a CC’s tendency to underrefer.

The host hospital 24-hour underreferral rate, as defined in this study, has many advantages as a measure of a CC’s tendency to underrefer children in need of prompt examination. It is an objective measurement. The data necessary to determine this rate are likely to be available from the host hospital’s computer systems. As illustrated in this study, it is likely that, for many pediatric CCs, their host hospital 24-hour underreferral rates would reasonably approximate their more-global 24-hour underreferral rates. These institutions could determine their host hospital 24-hour underreferral rates at regular intervals with minimal effort after initial configuration. The effects of protocol modifications or training initiatives on a CC’s tendency to underrefer could then be monitored and analyzed. Finally, at least for the CC studied, the month-to-month variability of the host hospital 24-hour underreferral rate was minimal, which supports its value as an ongoing monitor of a CC’s tendency to underrefer.

A pediatric CC is unlikely to ever achieve a 24-hour underreferral rate of 0%. Some children’s conditions are stable in all respects at the time of the call but deteriorate within hours after the call. With respect to that point, it is not surprising that two thirds of the underreferrals in this study involved patients with gastroenteritis, croup, asthma, or bronchiolitis, illnesses that are well known for their variable and unpredictable courses. Nevertheless, this study has demonstrated that a 24-hour underreferral rate of 5.95 ± 2.75% can be achieved by a high-volume pediatric CC. Determining whether a lower 24-hour underreferral rate can be achieved will require additional study.

Neither the host hospital 24-hour underreferral rate nor the more-global 24-hour underreferral rate should be viewed as an absolute measure of CC safety. For example, the rates would fail to identify children admitted 25 hours after inappropriate CC advice or children provided with homecare advice who visited an ED soon after receiving homecare advice, received urgent treatment, and were discharged. Nevertheless, these measures of CC safety provide an automated method of serially monitoring a CC’s safety.

The call another provider rate of 20.7% reported in this study is high, compared with similar rates reported in the literature.7 The protocols used in this study were designed specifically to triage acutely ill or injured, previously well children. Accordingly, all protocols contain questions to identify children with chronic medical conditions and those whose acute medical problems have been the subject of aggressive medical evaluation or treatment before the call. Such calls are screened for highly urgent conditions and then referred to the appropriate primary care provider or specialist. Historically, more than two thirds of all calls referred to speak with another provider by the Children’s Healthcare of Atlanta CC were referred because of those 2 questions. This philosophy almost certainly reduces underreferrals by the CC.

Review of individual calls that led to an underreferral was quite productive. For example, it became evident that several patients with croup who were admitted within 24 hours had not been prescribed steroids by their pediatricians the morning after the call. This information led to a change in the follow-up instructions to parents of children with croup. In another example, the child admitted with staphylococcal scalded skin syndrome presented with scalp tenderness, minimal erythema, and no fever. The nonfever arm of the rash protocol was modified to identify such a case in the future. Review of calls for patients who were admitted within 24 hours and were referred to a higher level of care by the CC allowed the reviewers to provide substantial positive feedback and medical follow-up information to the CC nurses.

This study has some limitations. The study was approved for and conducted during the winter months of October through March. Seasonal influences could have affected the underreferral rates determined here.

Sampling for admissions to a non–Children’s Healthcare of Atlanta facility used the method of convenience sampling. It is possible, for example, that admissions that received a nonurgent disposition occurred with greater frequency among noncontacts than among contacts completed during sampling. Other biases are always possible with a convenience sample. An additional limitation of telephone sampling, in this study and other studies, is determining accurately the timing of an admission to a facility without access to the medical charts. A limitation of our method of identifying admissions to the host hospital was the reliance on matching birth dates and last names. This undoubtedly resulted in missed admissions within 24 hours. However, we do not think that underreferrals would be disproportionately missed with this method.

Comparison with other CCs might be hampered by differences in the categorization of calls among CCs. For example, cases with significant chronic medical conditions are managed by Children’s Healthcare of Atlanta CC, but they are managed very conservatively. Most but not all of these calls are returned to a pediatrician or specialist for triage after emergency and highly urgent
situations are ruled out. These calls were included in this study. Other CCs might exclude such calls from statistical analysis. As another example, the calls that were categorized as information only by Children’s Healthcare of Atlanta CC and were eliminated from additional consideration might be accounted for in a different manner by other CCs.

Finally, whether an admission was or was not related to the illness or injury that prompted the call was sometimes subjective. However, as noted above, few calls were eliminated for this reason. Simply assuming that all admissions are related to the illness or injury that prompted the call makes the host hospital 24-hour underreferral rate easier to determine without significantly compromising data integrity (at least for the CC studied). Despite its limitations, we think that this study demonstrates a method that could be used by other CCs to monitor underreferral trends and to identify both opportunities for improvement and opportunities to provide positive feedback to CC nurses.

REFERENCES

WILL BRITAIN BE SIDELINED BY THE RISE OF ASIAN SCIENCE?

“When UK stem-cell researcher Stephen Minger embarked on a tour of stem-cell laboratories in China, Korea, and Singapore in 2004 with a group of western scientists, he soon lost any preconceptions he might have held about scientific facilities in Asia. In Shanghai especially, he admits to being ‘absolutely flabbergasted’ by the gleaming labs, high-tech equipment, and vast numbers of researchers. ‘Every lab we went to just got better and better.’ . . . China, India, and South Korea have seen major rises in scientific investment in the past few years. China’s research and development (R&D) spending has risen by 20% every year since 1999, and last year the country spent £4.7 billion compared with the UK’s £3.2 billion. South Korea meanwhile, spent 3% of its GDP (gross domestic product) on R&D in 2005, compared with the UK’s 1.86%. Between 2005 and 2006, the Indian government’s science spending rose by 24% to US$4.5 billion. The country is positioning itself as an R&D hub, with its pharmaceuticals market worth $8.2 billion—a sixth of the global market. Another indicator of the change in fortunes in Asian countries is the phenomenon of a ‘brain gain,’ a reverse of the brain drain of the past few decades that saw talented individuals emigrate west.”

Shetty P. Lancet. March 10, 2007

Editor’s Note: Research budgets in the United States are being cut or flat-lined in 2007–2008. Noted by JFL, MD
Intensive Home Visiting Is Associated With Decreased Risk of Infant Death

Edward F. Donovan, MD, Robert T. Ammerman, PhD, John Besl, MS, Harry Atherton, MS, Jane C. Khoury, PhD, Mekibib Altaye, PhD, Frank W. Putnam, MD, Judith B. Van Ginkel, PhD

Cincinnati Children’s Hospital Medical Center, Cincinnati, Ohio; Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, Ohio

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

ABSTRACT

OBJECTIVE. The goal was to test the hypothesis that participation in a community-based home-visiting program is associated with a decreased risk of infant death.

METHODS. A retrospective, case-control design was used to compare the risk of infant death among participants in Cincinnati’s Every Child Succeeds program and control subjects matched for gestational age at birth, previous pregnancy loss, marital status, and maternal age. The likelihood of infant death, adjusted for level of prenatal care, maternal smoking, maternal education, race, and age, was determined with multivariate logistic regression. The interaction between race and program participation and the effect of home visiting on the risk of preterm birth were explored.

RESULTS. Infants whose families did not receive home visiting (n = 4995) were 2.5 times more likely to die in infancy compared with infants whose families received home visiting (n = 1665). Black infants were at least as likely to benefit from home visiting as were nonblack infants. No effect of program participation on the risk of preterm birth was observed.

CONCLUSION. The current study is consistent with the hypothesis that intensive home visiting reduces the risk of infant death.
THE AMERICAN ACADEMY of Pediatrics encourages pediatricians to "support referral of high-risk parents to home-visitation programs as early as possible, ideally before or at the time of the prenatal visit to the pediatrician."1 Prenatal and infancy home visitation seeks to optimize pregnancy outcomes and child development through family education, training, and social support. Programs often target mothers and children at greater risk for adverse outcomes.

In a series of controlled trials, Olds et al2,3 found that home visiting by nurses reduced smoking during pregnancy, decreased preterm birth rates for smokers, increased birth weights among adolescent mothers, and decreased rates of child abuse and accidental injuries in children. Studying the effects of home visitation on infant mortality rates requires a large sample size, which may be available only in larger programs implemented in community settings. Examination of such programs is valuable in its own right, given the challenges of implementing programs originally developed in the setting of a controlled trial and given the need to better understand the effects of larger programs taken to scale.4

The purpose of the current study was to examine the impact of a large-scale, community-based, home visitation program. We tested the hypothesis that participation in greater Cincinnati’s Every Child Succeeds (ECS) program is associated with a decreased risk of infant death.

DESCRIPTION OF ECS PROGRAM

ECS is a community-based, home visitation program for first-time mothers and their children. ECS uses 2 national models of home visitation, namely, Nurse-Family Partnership (NFP)4 and Healthy Families America (HFA).5 Eligible mothers have ≥1 of 4 risk characteristics, that is, unmarried, inadequate income (up to 300% of poverty level, receipt of Medicaid, or reported concerns about finances), <18 years of age, or suboptimal prenatal care. Women are enrolled either during pregnancy (before 28 weeks for NFP) or before their child reaches 3 months of age (HFA only). Regular home visits are provided by social workers, child development specialists or related professionals (82%), trained nurses (12%), or paraprofessionals (6%). Home visits are made until the child reaches 2 years (NFP) or 3 years (HFA) of age, starting with weekly or more-frequent visits and tapering to fewer visits as the child ages. The goals of home visitation, as provided by ECS, are (1) to improve pregnancy outcomes through nutrition education and substance use reduction, (2) to support parents in providing children with a safe, nurturing, and stimulating home environment, (3) to optimize child health and development, (4) to link families to health care and other needed services, and (5) to promote economic self-sufficiency.

STUDY DESIGN AND METHODS

The study and use of Ohio birth certificates were approved by the Cincinnati Children’s Hospital Medical Center and the Ohio Department of Health institutional review boards. A retrospective, case-control design was used to examine the impact of home visiting on the likelihood of death before 1 year of age among infants enrolled in the ECS program. Study infants were limited to those residing in the ECS service area, including infants born to mothers residing in 1 of 4 southwestern Ohio counties and born between January 1, 2000, and December 31, 2001, and infants born to mothers residing in 1 of 3 northern Kentucky counties and born between January 1, 2000, and December 31, 2002. Infant deaths were limited to individuals who were residing in Ohio or Kentucky at the time of death.

An intent-to-treat approach was used, in which all mothers enrolled during the study periods who had ≥1 home visit were included in the sample. ECS provided records of all 2308 eligible and enrolled women and their infants born during the study periods. Among potentially eligible ECS participants, 643 were excluded from the analysis data set because no matching birth certificate was identified (n = 615) or the birth occurred in another state (n = 28) (Fig 1). Because participants could be included only if the birth and death certificates for those specific infants could be identified, we used the electronic ECS database to compare selected characteristics that increase the risk of infant death between participants included in the study data set and excluded ECS participants. There were more married mothers among the excluded participants, but no other differences were observed (Table 1).

ECS records were linked to Ohio or Kentucky birth certificates by using common data fields. Because Ohio does not provide the mother’s Social Security number on its electronic birth certificates, 4 fields in the ECS data

![FIGURE 1](Enrollment flowchart)
set were potentially used for linkage with the appropriate Ohio birth certificate, namely, mother’s name, mother’s date of birth, child’s name, and child’s date of birth. Agreement in ≥2 fields was required for qualification as a match. For Kentucky residents, a hierarchical linking algorithm that used the mother’s Social Security number in the first step was invoked. Because women might have had >1 birth during the study period, ECS records were linked to Kentucky births by using the Social Security number and birth year.

For linkage of Ohio cases to birth certificates, any combination of 2 or 3 fields was subjected to additional review. Additional variables were used for verification purposes. Some potential matches were rejected when, for instance, the birth certificate recorded a male birth but the ECS child was recorded as female. The number of previous live births is a birth certificate data item that was used as a check for finding the correct birth certificate for each ECS participant. Because ECS targets first-time mothers, a birth certificate identifying a second or higher-order birth was considered suspect. Whenever a potential match was rejected, the record was restored to both the ECS and state birth certificate data sets, so that it would be available for matching in subsequent combinations of fields.

For the linked ECS-Kentucky births, birth certificates were linked to electronic Kentucky or Ohio death certificates by using the birth certificate number that is included on the death certificate. If a child is born in Kentucky but dies elsewhere, then the birth certificate number field carries the number assigned by the birth state. For linked ECS-Ohio births, the electronic Ohio birth certificates were merged with an electronic linked birth-infant death file provided for infant deaths that occurred in 2000, 2001, or 2002. Because a linked birth-infant death file was not available for 2002 Ohio births, these could not be included in the analyses.

Three control subjects for each ECS study subject were selected randomly from among births to first-time mothers that occurred in the same time period to mothers residing in the ECS service area who were either <18 years of age or unmarried at the time of birth. Because ECS enrolls participants either during pregnancy or after birth, 2 different algorithms were used to select control subjects. Control subjects for infants enrolled during the pregnancy were selected randomly, without replacement, from among pregnancies that had not been delivered by the week of gestation of enrollment of the ECS subject. For example, if ECS enrolled a mother when the fetal gestational age was 26 weeks, then 3 control subjects were selected from among eligible infants not enrolled in ECS whose pregnancy duration was ≥26 weeks. For each ECS subject enrolled after birth, 3 control subjects were selected randomly, without replacement, by using 2 criteria, such that control subjects had both the same gestational age at birth and were alive at the same postmenstrual age as the postmenstrual age of enrollment of the index ECS subject. For ECS case subjects enrolled after birth because of poor prenatal care, the Kotelchuck index was used as an additional criterion for selecting control subjects. Therefore, control subjects for these ECS subjects were matched for gestational age at birth, postmenstrual age at ECS enrollment, and 1 of the 4 Kotelchuck index prenatal care categories (inadequate, intermediate, adequate, or adequate plus).6

## ANALYSES

The aforementioned selection of ECS and control subjects and all data analyses were performed by using SAS 9.1 (SAS Institute, Cary, NC). Adjusted comparisons of the likelihood of death before 1 year of age between ECS participants and control subjects were made by using multivariate logistic regression. Infants with trisomy 18, trisomy 13, anencephaly, or renal agenesis were not included in the study data set, because such anomalies are considered lethal and nonpreventable.

The primary independent variable was participation in ECS. Potential adjustment variables were selected from those available in the birth certificate file that are known to be associated with infant death, including mother’s race (black versus not black, reflecting the race/ethnicity distribution for greater Cincinnati), maternal age <19 years, suboptimal prenatal care,6 single marital status at the time of birth (as reported by the mother), maternal smoking during the pregnancy, suboptimal maternal education (>1 year behind the expected grade level if <19 years of age or not a high school graduate if ≥19 years of age), and the race/ethnicity distribution for greater Cincinnati. Some potential matches were rejected when, for instance, the birth certificate recorded a male birth but the ECS subject was recorded as female. The number of previous live births is a birth certificate data item that was used as a check for finding the correct birth certificate for each ECS participant. Because ECS targets first-time mothers, a birth certificate identifying a second or higher-order birth was considered suspect. Whenever a potential match was rejected, the record was restored to both the ECS and state birth certificate data sets, so that it would be available for matching in subsequent combinations of fields.

## TABLE 1

Comparison of ECS Participants With Linked Birth Certificates and ECS Participants Without Linked Birth Certificates or Not Born in Ohio or Kentucky

<table>
<thead>
<tr>
<th></th>
<th>ECS Participants Included in Study Sample</th>
<th>ECS Participants Excluded From Study Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>1665</td>
<td>643</td>
</tr>
<tr>
<td>Age of mother, mean ± SD, y</td>
<td>19 ± 4</td>
<td>20 ± 4</td>
</tr>
<tr>
<td>Gestational age, mean ± SD, wk</td>
<td>39 ± 3</td>
<td>38 ± 5</td>
</tr>
<tr>
<td>School completed, mean ± SD, y</td>
<td>11 ± 2</td>
<td>11 ± 2</td>
</tr>
<tr>
<td>Not married, %</td>
<td>96</td>
<td>83</td>
</tr>
<tr>
<td>Black, %</td>
<td>33</td>
<td>35</td>
</tr>
</tbody>
</table>

* At ECS eligibility determination.
previous pregnancy, and product of a multifetal gestation.

Although congenital malformations are associated with infant death, no deaths among infants in the study data set involved infants with congenital malformations. Mothers with absent data on smoking were coded as smokers. This had little effect, in the regression models, on the coefficients for either the effect of ECS or the effect of smoking on the likelihood of infant death. A final set of variables to be included in the initial regression analyses was determined on the basis of the total number of observed deaths, the distribution of each independent variable among study infants, and the strength of the previously reported association between the variable and the likelihood of infant death.

For ECS case subjects enrolled before birth, ECS might have influenced the likelihood of receipt of prenatal care. To evaluate the relationship between ECS participation and adequacy of prenatal care, we compared the parameter estimate for ECS in the same model with the parameter estimate for ECS in the final model excluding the prenatal care variable.

Because infant mortality rates in the United States are greater among black infants, compared with white infants, we used a parsimonious, multivariate, logistic regression model to calculate predicted probabilities for comparisons of risk of infant death among black infants participating in ECS, black control subjects, nonblack infants participating in ECS, and nonblack control subjects. Because one possible benefit of prenatal enrollment would be to reduce the risk of preterm birth, we used the same independent variables in a linear regression model of the association between prenatal ECS enrollment and gestational age at birth, including as observations only the ECS subjects and associated control subjects enrolled before delivery.

RESULTS Among 1665 ECS case subjects included in the analyses, 715 (43%) were enrolled in ECS before birth, 950 (57%) were enrolled after birth, 1386 (83%) received home-visiting services based on the HFA model, and 279 (17%) received home-visiting services based on the NFP model. The mean ± SD gestational ages of ECS subjects enrolled prenatally in the HFA and NFP subsections of the ECS program were 29 ± 8 and 20 ± 7 weeks, respectively. ECS participants who were enrolled before birth and survived to 1 year of age had a mean of 6.9 home visits before birth and 18.3 visits between birth and 1 year of age. ECS participants who were enrolled after birth and survived to 1 year of age had a mean of 14.3 home visits before the age of 1 year. Characteristics of infants enrolled in ECS before birth, infants enrolled after birth, and their respective control subjects are displayed in Table 2. These characteristics reflect efforts by ECS to target high-risk women having their first child, that is, adolescent, unmarried, low income, or suboptimal education. The rate of maternal smoking among study participants was ~3 times that reported for the overall population of mothers in the United States.

On the basis of the number of deaths in the study sample, the final model to test the effect of ECS on the likelihood of death included 7 independent variables. Three models are displayed in Table 3. Model 1 includes the 9 proposed independent variables.

Because there were few infants of multifetal gestations and because the coefficient for the ECS variable did not change when this variable was removed, the multifetal gestation variable was not included in model 2. Because there were few married women and because the coefficient for the ECS variable did not change when this variable was removed, neither multifetal gestation nor marital status was included in model 3. The association between enrollment in ECS and the likelihood of infant death was the same in all models tested. After.

### Table 2

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Prenatal</th>
<th>Postnatal</th>
<th>Control</th>
<th>Prenatal</th>
<th>Postnatal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Death, % (n)</td>
<td>0.7 (5)</td>
<td>2.0 (45)</td>
<td>0.2 (2)</td>
<td>0.3 (8)</td>
<td></td>
</tr>
<tr>
<td>Black, %</td>
<td>28</td>
<td>30</td>
<td>37</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Suboptimal prenatal care, %a</td>
<td>21</td>
<td>31</td>
<td>30</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>Multifetal gestation, % (n)</td>
<td>0.8 (6)</td>
<td>1.0 (22)</td>
<td>0.4 (4)</td>
<td>1.1 (31)</td>
<td></td>
</tr>
<tr>
<td>Smoking, %b</td>
<td>35</td>
<td>27</td>
<td>30</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Suboptimal education, %</td>
<td>26</td>
<td>19</td>
<td>25</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Unmarried, %</td>
<td>96</td>
<td>97</td>
<td>97</td>
<td>98</td>
<td></td>
</tr>
<tr>
<td>Mother &lt;19 y of age, %</td>
<td>51</td>
<td>29</td>
<td>44</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>Any previous fetal death, %</td>
<td>13</td>
<td>15</td>
<td>11</td>
<td>14</td>
<td></td>
</tr>
</tbody>
</table>

a Suboptimal prenatal care was defined as intermediate or inadequate with the Kotelchuck index, compared with adequate or adequate plus.

b Mothers with absent data on smoking were coded as smokers, in part because this had little effect in the regression models on the coefficients for either the effect of ECS or the effect of smoking on the likelihood of infant death.

### Table 3

<table>
<thead>
<tr>
<th>Variable</th>
<th>Odds Ratio (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECS</td>
<td>0.4 (0.2–0.9)</td>
</tr>
<tr>
<td>Black</td>
<td>2.7 (1.6–4.7)</td>
</tr>
<tr>
<td>Suboptimal prenatal care</td>
<td>3.2 (1.9–5.4)</td>
</tr>
<tr>
<td>Multifetal gestation</td>
<td>7.3 (2.4–21.9)</td>
</tr>
<tr>
<td>Maternal smoking</td>
<td>0.8 (0.4–1.7)</td>
</tr>
<tr>
<td>Suboptimal education</td>
<td>0.9 (0.5–1.8)</td>
</tr>
<tr>
<td>Married</td>
<td>1.7 (0.4–7.3)</td>
</tr>
<tr>
<td>Mother &lt;19 y of age</td>
<td>1.1 (0.7–2.0)</td>
</tr>
<tr>
<td>Any previous fetal death</td>
<td>2.2 (1.2–4.1)</td>
</tr>
</tbody>
</table>

ECS indicates enrolled in the ECS program, compared with nonenrolled control subjects; black, black race, compared with all other races; suboptimal prenatal care, Kotelchuck index result of intermediate or inadequate, compared with adequate or adequate plus; multifetal gestation, twin or triplet; maternal smoking, mother reported any smoking during pregnancy; suboptimal education, years of education >1 year less than expected for age; married, married at delivery of child, compared with single, divorced, or widowed.
controlling for race, prenatal care, maternal smoking, maternal education, and maternal age, enrollment in ECS was associated with a 60% decrease in the likelihood of infant death. When inadequate prenatal care was removed as an independent variable from model 3, the parameter estimate for ECS decreased slightly, from 0.41 (Table 3) to 0.38 (model not shown), consistent with the possibility that ECS increased the adequacy of prenatal care.

The majority of deaths among study infants involved the ECS subjects and their respective control subjects enrolled before birth (Table 2). We were unable to identify different predictors for infants enrolled in ECS before birth and those enrolled after birth. Table 4 displays the results of multivariate linear regression analysis evaluating the association between ECS enrollment and gestational age at birth for ECS subjects and control subjects identified before delivery. No influence of ECS enrollment on gestational age at birth was observed. The difference between the adjusted predicted probabilities of death for black infants participating in ECS and black control subjects tended to be larger than the corresponding difference for nonblack infants (Fig 2).

DISCUSSION

In 2002, the US infant mortality rate ranked 24th worst among nations reporting to the World Health Organization. The health status of US residents overall and the health status of many US subgroups have improved remarkably in nearly all measures in recent decades. Many individuals enjoy better health, but large disparities persist among racial/ethnic, socioeconomic, and geographic subgroups. It is these disparities that, in very large part, explain the persistently low US ranking in international comparisons. For example, if the US infant mortality rate for black infants (14.4 deaths for every 1000 live births in 2002) were similar to that for white infants (5.8 deaths for every 1000 live births), then the US ranking among reporting United Nations member countries would increase from 24th to 7th. We report findings that support an association between intensive home visiting for high-risk, first-time, pregnant mothers and reduced likelihood of infant death. Nonparticipants were 2.5 times more likely to die in infancy, compared with those enrolled in ECS. These findings expand our understanding of the impact of home visitation and should guide future interventions designed to decrease infant mortality rates.

Causes of infant death can be categorized as possibly preventable or not preventable. The much higher infant mortality rate for US black infants, compared with white infants, suggests that some black deaths may be preventable. In the current study, the differences between black participants in ECS and nonparticipants tended to be much larger than differences observed for nonblack infants. Our analyses showed that adequacy of prenatal care had the strongest association with the likelihood of infant death. Black mothers enrolled in ECS before birth were more likely to receive adequate prenatal care, compared with control subjects (74% vs 60%), which might have been one of the determinants of the lower infant mortality rate seen in that group. Adequacy of prenatal care, as measured with the Kotelchuck index, includes both pregnancy month at initiation of prenatal care and number of prenatal visits. Because ECS could have no impact on initiation of prenatal care, we evaluated the study model by using month of initiation of prenatal care, rather than the entire Kotelchuck index, as a covariate. Therefore, any effect of the ECS program on the number of prenatal visits was embedded in the ECS variable. No change in the effect of ECS on the likelihood of infant death was observed (odds ratio: 0.43 vs 0.41).

We cannot make inferences about causation because of the retrospective, case-control design used in this study. However, it is encouraging that the largest association between ECS participation and reduced infant mortality rate was seen for black infants. It is not evident specifically how participation in ECS might result in reduced risk of infant death. The number of infant

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\beta$ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECS</td>
<td>0.19 ($-0.07$ to $0.46$)</td>
</tr>
<tr>
<td>Black</td>
<td>$-0.65$ ($-0.91$ to $-0.39$)</td>
</tr>
<tr>
<td>Suboptimal prenatal care$^a$</td>
<td>0.55 ($0.30$ to $0.81$)</td>
</tr>
<tr>
<td>Maternal smoking$^b$</td>
<td>$-0.09$ ($-0.18$ to $0.35$)</td>
</tr>
<tr>
<td>Suboptimal education</td>
<td>$-0.02$ ($-0.27$ to $0.30$)</td>
</tr>
<tr>
<td>Mother &lt;19 y of age</td>
<td>0.08 ($-0.16$ to $0.32$)</td>
</tr>
</tbody>
</table>

$^a$ Suboptimal prenatal care was defined as intermediate or inadequate with the Kotelchuck index, compared with adequate or adequate plus.

$^b$ Mothers with absent data on smoking were coded as smokers, in part because this had little effect in the regression models on the coefficients for either the effect of ECS or the effect of smoking on the likelihood of infant death.
deaths in the current study does not allow careful analysis of age of death and cause of death. The largest contributors to infant death in the United States are preterm birth, complications of labor and delivery, congenital anomalies, sudden infant death syndrome (SIDS), infections, and injury. Except for congenital anomalies, home visiting theoretically could influence any of these death mechanisms. For women enrolled in ECS before delivery, improving the amount and content of prenatal care may decrease the risk of preterm birth, perinatal complications, and SIDS through maternal education and identification/amelioration of barriers to optimal care. For injuries (intentional and unintentional), infection, SIDS, and lasting complications of perinatal disease, home visiting may reduce the risk of death during infancy through similar mechanisms.

In well-designed randomized trials, home visiting has been shown to have important short-term and long-term benefits. In a randomized trial that compared provision of one half of indicated prenatal visits in the home by a nurse specialist and usual prenatal care for 173 high-risk pregnant women (94% black), the mean number of prenatal visits was increased and the home visit group had fewer preterm births and fewer combined fetal/infant deaths. Randomized trials of combined prenatal and infancy home visiting by nurses demonstrated decreases in reported and substantiated child abuse and neglect rates, fewer emergency department visits, fewer physician visits for treatment of accidents and poisonings, and healthier subsequent pregnancies. The average birth weight for mothers <17 years of age was 395 g greater than that for a comparison group with no home visits, and preterm delivery among smoking mothers was decreased by 75% with home visiting. In an uncontrolled study, home visiting among a cohort of high-risk women was associated with increased use of prenatal care. A prospective observational study suggested that pregnant, high-risk women who had 5 to 9 home visits during pregnancy had higher average birth weights, had fewer low birth weight infants, and were more likely to breastfeed. These studies suggest that home visiting, particularly during pregnancy, may help to reduce the risk of infant death by decreasing preterm birth, low birth weight, and child abuse/neglect. In the current study, home visiting was associated with better prenatal care, which was, in turn, associated strongly with decreased risk of infant death. Because less than one half of ECS families were enrolled before birth, we were unable to determine the extent to which prenatal home visits might have improved the adequacy of prenatal care.

Translating findings from randomized trials to the real world can be problematic. evaluated the effect of a statewide nurse home-visiting program on the risk of infant death by using Oklahoma birth and death certificates. Children of first-time mothers were compared with eligible nonparticipants. For high-risk pregnancies, participation was associated with decreased likelihood of birth at gestational age of <37 weeks and decreased infant mortality rates, with odds ratios very similar to those found in our study. Observed benefits of home visiting may depend on the qualifications of the home visitors, with larger effects being seen with nurses, compared with other types of home visitors. It is encouraging that use of home visitors with a variety of qualifications was associated with reduced infant mortality rates in the Cincinnati region. The current study has inadequate sample size for determination of the independent effects of the type of home visitor or the independent effects of the various ECS program components.

Confidence in inferences that can be made on the basis of retrospective case-control studies is limited. It is particularly difficult to address possible bias in the selection of control subjects. Inherent in the ECS enrollment process is the requirement that women initially agree to participate and the possibility that these women are inherently different from women who do not or cannot enroll. To help address this issue, we conducted a time-limited comparison of mothers who agreed to participate and those who refused to participate. This comparison of 625 mothers (agreed, n = 315; refused, n = 310) occurred over a 6-month period at birth hospitals in which all eligible mothers were approached and offered home visitation services. Mothers were compared with respect to 16 psychosocial risk factors (eg, social isolation, unemployment, and housing instability) used by many HFA programs. No significant differences between groups with respect to these risk factors or race were observed (P > .05), except that enrolled mothers were slightly younger than those who refused (18.7 vs 19.4 years).

Our study findings are consistent with the findings from randomized, controlled trials and suggest that home visiting reduces the risk of infant death. To identify and to characterize ECS participants, we used an administrative data set that was designed for management and evaluation of the ECS program, rather than scientifically rigorous research. We were unable to link a substantial portion (25%) of ECS participants to their respective birth and death certificates. It is possible that those ECS participants who were excluded from the study were systematically different from those who were included. However, we were unable to identify a rational explanation for such bias, and our comparison of risk characteristics for the included and excluded case subjects demonstrated no differences that would bias the study in a direction that favored ECS (Table 1). Although included ECS case subjects were less likely to be married (4% vs 17%), this would most likely bias the study in the direction of more-favorable outcomes among control infants. Moreover, marital status in our regression mod-
els had a weak association with the likelihood of infant death.

Infant death may be considered the “tip of the iceberg” in which the children of families at risk experience suboptimal care, poor health outcomes, and the possibility of lifelong disability; some die before their first birthday. The current study is consistent with the hypothesis that intensive home visiting reduces the risk of infant death.

ACKNOWLEDGMENTS
This work was funded in part by the Children’s Hospital Research Foundation (Cincinnati, OH) and the Maternal and Child Health Bureau (grant R40MC06632). ECS acknowledges the support of its founding partners: the United Way of Greater Cincinnati, Cincinnati-Hamilton County Community Action Agency, and Cincinnati Children’s Hospital Medical Center.

REFERENCES

RANDOMIZED CLINICAL TRIALS

Pediatrics requires investigators to register their clinical trials in a public trials registry. The members of the International Committee of Medical Journal Editors (ICMJE) will consider most clinical trials for publication only if they have been registered (see N Engl J Med. 2004;351:1250–1251). Current information on requirements and appropriate registries is available at www.icmje.org/faq.pdf.
ARTICLE

Endothelial Function in Newborn Infants Is Related to Folate Levels and Birth Weight

Helena Martin, MD, PhD, Bo Lindblad, MD, PhD, Mikael Norman, MD, PhD

Department of Woman and Child Health, Division of International Health, Department of Public Health Sciences, and Department of Clinical Science, Intervention, and Technology, Karolinska Institute, Stockholm, Sweden

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

ABSTRACT

OBJECTIVE. Low maternal folate levels during pregnancy correlate with low birth weight, a perinatal risk factor for later cardiovascular disease. We studied relationships between red blood cell folate levels, birth weight, and vascular endothelial function (a key factor in the early pathophysiologic processes of heart disease) in newborn infants.

METHODS. We included 82 infants (30 low birth weight) and their mothers. A laser Doppler technique was used to measure skin perfusion during transdermal iontophoresis of acetylcholine (an endothelium-dependent vasodilator). Red blood cell folate, vitamin B₁₂, and homocysteine levels were determined.

RESULTS. The perfusion response to acetylcholine was lower in low birth weight infants than in normal birth weight control subjects (mean: 35 vs 76 perfusion units). The neonatal acetylcholine response correlated with red blood cell folate levels in both infants and their mothers. The folate levels of low birth weight and control infants did not differ significantly (mean: 1603 vs 1795 nmol/L), but mothers of low birth weight infants had lower folate levels than did mothers of control infants (mean: 805 vs 1109 nmol/L). In multivariate analysis, low birth weight and red blood cell folate levels contributed independently to endothelial function in newborn infants. The levels of vitamin B₁₂ and homocysteine were similar in the 2 groups and did not correlate with endothelial function.

CONCLUSION. The data presented here provide the first evidence for a relationship between folate levels and vascular endothelial function in newborn infants.
Dysfunction of the vascular endothelium precedes formation of atheromatous plaques and is linked to metabolic and sympathetic abnormalities. The degree of endothelial dysfunction is related to the severity and prognosis of heart disease. Evaluation of endothelial function thus permits assessment of cardiovascular risk in healthy subjects and study of pathogenetic mechanisms. Already as newborn infants, subjects with low birth weight (LBW) show an impairment in endothelium-dependent vasodilation. Endothelial dysfunction also characterizes LBW children and adults. At present, the causes of endothelial damage to human fetuses are not known. Experimental data show that nutritional imbalances in pregnancy cause lasting endothelial dysfunction in the offspring.

Low folate and high homocysteine levels, in combination or independently, have been shown to be risk factors for endothelial dysfunction and cardiovascular disease. Their metabolism is inversely correlated. The molecular mechanisms proposed include oxidative inactivation and reduced synthesis of the endothelium-derived vasodilator nitric oxide (Fig 1). During pregnancy, low folate and high homocysteine levels are associated with maternal vascular complications, for example, spontaneous abortion, preeclampsia, and placental abruption. More recently, low folate and high homocysteine levels were found to be associated with LBW. The associations with lower birth weight cannot be attributed to folate-related genetic polymorphisms alone or blocked placental transport and binding of folate but are ascribed mainly to deficient dietary intakes. In randomized trials, high but not low doses of folic acid during pregnancy reduced the prevalence of LBW.

We measured endothelium-dependent vasodilation in LBW and normal birth weight (NBW) infants born at term and their mothers. We hypothesized that impairment in vascular endothelial function would be associated with low folate/high homocysteine levels. Because vitamin B12 deficiency in pregnancy may cause the accumulation in tissue of homocysteine and lower birth weight, the levels of vitamin B12 were also determined.

Methods

Subjects

We studied 82 newborn infants and their mothers, for a total of 164 subjects (Table 1). Parental informed consent was obtained, and the study was approved by the regional ethics committee of Karolinska Institute. All mothers were healthy nonsmokers who used no medications or special diets. Subjects with manifest (blood pressure of >140/90 mm Hg) or borderline (diastolic blood pressure of >85 mm Hg) hypertension or with insulin-dependent diabetes mellitus or glucose intolerance during the index pregnancy were excluded, as were multiple pregnancies and infants with congenital infections, chromosomal disorders, malformations, or neonatal asphyxia. Gestational age was determined through ultrasonography, and only infants born after 35 weeks of gestation were included.

Using these criteria, we identified 41 consecutive cases of LBW infants born at Danderyd Hospital in Stockholm between September 2003 and September 2004. LBW was defined as a birth weight less than the mean – 2 SD (ie, below the gender-specific 2.5th percentile for gestational age), according to Swedish reference data for normal fetal growth. Among the eligible infants, 9 were lost because of lack of investigator time and 2 because their parents declined to participate in the study. Thirty case subjects and 52 NBW control subjects were included. The birth weights of the 2 groups did not overlap (Table 1).

In the LBW group, prenatal fetometry using ultrasonography showed impaired growth for 16 of 30 fetuses. An obstetrical decision to deliver through cesarean section was made for 12 of 30 in the LBW group. The NBW control group was matched with respect to postnatal and gestational ages, gender, and mode of delivery.

Of women delivering a LBW infant, 3 had gestational hypertension and 4 had preeclampsia (gestational hypertension plus proteinuria of >1+). No mother who delivered a NBW infant had hypertension during pregnancy.

All infants were breastfed by their mothers in the maternity ward. Extra formula was provided to a larger proportion of LBW infants (11 of 30 infants), compared with NBW infants (7 of 52 infants; P < .05). On the day of examination, all infants had plasma glucose and hematocrit levels within the reference ranges (Table 1). Parents were interviewed concerning a family history of diabetes mellitus, myocardial infarction, stroke, hypertension, or hyperlipidemia among first-degree relatives.

Vascular Studies

Endothelial function was determined once, 3 to 4 days after delivery. The investigation was performed within 1 hour after feeding of the infants, while they were sleeping in the prone position. Mothers did not eat or drink for 2 hours before measurements. During the study, they sat comfortably in chairs, resting their forearms and hands on armrests at heart level.

A laser Doppler (LD) instrument (Periflux 4001; Perimed, Stockholm, Sweden) and a micropharmacology system were used to measure skin perfusion before and after transdermal delivery of acetylcholine, an endothelium-dependent vasodilator. The LD signal is proportional to the number and velocity of moving blood cells in the illuminated superficial skin microvessels and is expressed in perfusion units (PUs) of output voltage (1 PU = 10 mV). The combined drug-delivery and LD probe was fixed to the dorsal aspect of the hand with...
double-adhesive tape. The temperature of the LD probe facing the skin was standardized to 32°C. To study endothelium-dependent vasodilation, basal skin perfusion was recorded for 2 minutes, after which 2% acetylcholine chloride (Sigma-Aldrich, Steinheim, Germany) was transferred across the skin through iontophoresis (anodal current of 0.1 mA for 20 s, repeated 5 times at 60-s intervals). Basal perfusion and changes in response to acetylcholine were measured as the area under the curve. Details of the methods have been discussed elsewhere.6 Because of movement artifacts in infants, successful LD recordings were obtained for 78 of 82 subjects.

**Biochemical Analyses**

Blood samples were drawn from mothers and infants after vascular measurements. Plasma was separated within 30 minutes and stored at −20°C until analysis. Red blood cell (RBC) folate and vitamin B12 levels were determined with fluoroimmunoassays and plasma homocysteine levels with a fluorescence-polarization immunoassay.24,25 Maternal blood sampling and analyses were complete for all subjects. For healthy newborn infants, repeated blood sampling was not approved by the ethics committee. Because the metabolic screening program had first priority, sufficient sampling volumes and successful analyses of RBC folate and plasma homocysteine levels were obtained for 54 of 82 newborn infants.

**Statistical Analyses**

Values are means ± SEMs or numbers of subjects and proportions. Student’s t test, analysis of variance, or the χ² test was used to compare groups of data. Correlation coefficients were calculated and regression analyses were performed to evaluate contributions to differences in endothelial function (outcome). In these calculations, birth weight, levels of folate, homocysteine, and vitamin B12, family history of cardiovascular disease, gestational hypertension, mode of delivery, and neonatal hematocrit levels were included as covariates. Initially, group comparisons or linear regressions were performed for each covariate one by one. Covariates with P < .20 were entered into a multivariate regression model. Assess-
ments of perfusion responses to acetylcholine provocations in the 2 groups were made by using 2-factor analysis of variance for repeated measurements. We planned to include at least 30 case subjects and 30 control subjects, to detect a group difference in endothelium-dependent vasodilation of 0.7 SD at a significance level of .05 and power of 0.80.

RESULTS

Folate, Homocysteine, and Vitamin B12 Levels
The RBC folate levels in LBW and NBW infants were not significantly different (ie, 1603 ± 87 vs 1795 ± 72 nmol/L; P = .11). Mothers of LBW infants had lower RBC folate levels (805 ± 63 nmol/L) than did mothers of NBW infants (1109 ± 86 nmol/L; P = .02). We found no differences in plasma homocysteine levels between LBW and NBW infants (6.5 ± 0.4 vs 6.9 ± 0.3 μmol/L; P = .48) or between their mothers (9.7 ± 0.6 vs 9.1 ± 0.4 μmol/L; P = .40).

The vitamin B12 concentrations were 386 ± 36 pmol/L for LBW infants and 362 ± 72 pmol/L for NBW infants (P = .92). Maternal vitamin B12 levels were 222 ± 18 pmol/L for mothers of LBW infants and 191 ± 14 pmol/L for mothers of NBW infants (P = .20). There were significant correlations between folate levels in maternal and neonatal blood (r = 0.47; P < .001), between maternal and neonatal plasma homocysteine levels (r = 0.59; P < .0001), and between maternal and neonatal vitamin B12 levels (r = 0.79; P < .0001).

Endothelial Function and Birth Weight
Basal skin perfusion was 14 ± 0.7 PUs in LBW infants and 14 ± 0.8 PUs in NBW infants (P = .56). Acetylcholine-induced, endothelium-dependent vasodilation was significantly lower in LBW infants than in NBW infants. Peak perfusion induced by acetylcholine was 35 ± 3 PUs in LBW infants and 76 ± 5 PUs in NBW infants (P < .001) (Fig 2).

Mothers of LBW infants had lower basal skin perfusion (9.0 ± 1 PUs) than did mothers of NBW infants (12 ± 0.6 PUs; P = .002). Acetylcholine-induced, endothelium-dependent, peak perfusion was lower (with borderline significance) in mothers of LBW infants (94 ± 9 PUs) than in mothers of NBW infants (116 ± 8 PUs; P = .07) (Fig 2).

Endothelial Function in Relation to Folate Levels
In newborn infants, basal skin perfusion showed no correlation with the results of blood analyses. Endothelium-dependent, peak perfusion correlated with neonatal (r = 0.43; P = .003) and maternal (r = 0.36; P = .004) RBC folate levels (Fig 3). We found no associations between endothelial function and plasma homocysteine and vitamin B12 levels. In mothers, basal skin perfusion correlated with RBC folate levels (r = 0.25; P < .05). No associations between maternal endothelial function and RBC folate, plasma homocysteine, and vitamin B12 levels were noted.

Endothelial Function in Relation to Other Covariates
Endothelial function did not differ in infants with (20 of 82 infants) or without a positive family history of cardiovascular disease (P = .28). All infants of mothers with a history of gestational hypertension or preeclampsia (7 of 82 infants) were in the LBW group and had lower endothelium-dependent peak perfusion levels, com-

### TABLE 1 Characteristics of Study Groups

<table>
<thead>
<tr>
<th></th>
<th>LBW (N = 30)</th>
<th>NBW (N = 52)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Maternal data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, mean ± SEM, y</td>
<td>32.1 ± 0.9</td>
<td>32.9 ± 0.7</td>
<td>.47</td>
</tr>
<tr>
<td>Primipara, n (%)</td>
<td>23 (77)</td>
<td>23 (44)</td>
<td>.004</td>
</tr>
<tr>
<td>BMI, mean ± SEM, kg/m²</td>
<td>22.4 ± 0.5</td>
<td>23.1 ± 0.4</td>
<td>.34</td>
</tr>
<tr>
<td>Family history of CVD, n (%)</td>
<td>13 (43)</td>
<td>16 (31)</td>
<td>.25</td>
</tr>
<tr>
<td>Smoker, n</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td><strong>Pregnancy data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight gain, mean ± SEM, kg</td>
<td>10.7 ± 0.7</td>
<td>12.5 ± 0.7</td>
<td>.08</td>
</tr>
<tr>
<td>Hypertension, n (%)</td>
<td>7 (23)</td>
<td>0</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Length of gestation, mean ± SEM, wk</td>
<td>39 ± 0.2</td>
<td>39 ± 0.2</td>
<td>.52</td>
</tr>
<tr>
<td><strong>Infant data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boys, n (%)</td>
<td>14 (47)</td>
<td>25 (48)</td>
<td>.90</td>
</tr>
<tr>
<td>Birthweight, mean ± SEM, g</td>
<td>2394 ± 46</td>
<td>3571 ± 65</td>
<td></td>
</tr>
<tr>
<td>Birthweight deviation, mean ± SEM, %</td>
<td>-32 ± 0.8</td>
<td>1.1 ± 0.9</td>
<td></td>
</tr>
<tr>
<td>Birth length, mean ± SEM, cm</td>
<td>46.7 ± 0.4</td>
<td>50.1 ± 0.2</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Head circumference, mean ± SEM, cm</td>
<td>33.0 ± 0.2</td>
<td>35.4 ± 0.2</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Postnatal age, mean ± SEM, h</td>
<td>79 ± 1</td>
<td>77 ± 1</td>
<td>.11</td>
</tr>
<tr>
<td>Plasma glucose level, mean ± SEM, mmol/L</td>
<td>4.0 ± 0.2</td>
<td>4.1 ± 0.2</td>
<td>.67</td>
</tr>
<tr>
<td>Hematocrit level, mean ± SEM, %</td>
<td>55 ± 1.7</td>
<td>49 ± 1.1</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

Hypertension includes gestational hypertension or preeclampsia. CVD indicates cardiovascular disease.
pared with infants of mothers without such a history \((P = .02)\). Within the LBW group, endothelium-dependent peak perfusion did not differ in infants with or without a history of gestational hypertension or preeclampsia \((P = .48)\). Endothelial function did not correlate significantly with mode of delivery \((P = .08)\), but an inverse correlation between neonatal hematocrit levels and endothelium-dependent peak perfusion was found \((r = -0.30; P < .05)\). Maternal endothelial function did not correlate with a family history of cardiovascular disease or gestational hypertension.

**Multivariate Analysis of Endothelial Function**

Multivariate analysis with infant endothelial function as outcome and birth weight, gestational hypertension, mode of delivery, and infant RBC folate and hematocrit levels as independent risk factors showed that both LBW \((P < .001)\) and RBC folate levels \((P < .05)\) contributed independently to endothelium-dependent peak perfusion. Taken together, these risk factors could explain 41% of the estimated variance in neonatal endothelial function \((R^2 = 0.41; P < .001)\). Gestational hypertension \((P = .40)\), mode of delivery \((P = .18)\), and neonatal hematocrit levels \((P = .88)\) did not contribute to neonatal endothelium-dependent peak perfusion.

**DISCUSSION**

This is the first human study demonstrating a relationship between low maternal/neonatal folate levels and microvascular endothelial dysfunction in newborn infants. The observation is important. First, it provides insight into intrauterine mechanisms that can adversely affect the developing human vascular system. Second, neonatal endothelial damage may extend long after the perinatal period, ultimately affecting the risk of cardiovascular disease in adult life. Finally, trials with folate supplementation, at least at high doses, in pregnancy have shown a reduced risk of LBW and, as indicated by animal data, folate supplementation during pregnancy may abolish vascular endothelial dysfunction in the offspring.

One of the earliest signs of susceptibility to atherosclerosis is endothelial dysfunction, which precedes structural vascular lesions by decades. We and others previously found impaired endothelium-dependent vasodilation in infants, children, and young adults born small at term. Because systemic endothelial dysfunction in peripheral and coronary arteries is related to cardiovascular end points, these findings may facilitate understanding of the epidemiologic associations between LBW and heart disease. Although the causal pathway in early human life is still unclear, findings in animals show that permanent vascular endothelial dysfunction ensues in the adult offspring of diet-challenged pregnancies.

Mothers of LBW infants showed lower basal skin perfusion and a trend toward lower endothelium-dependent peak perfusion responses. Therefore, the moth-
ers of LBW infants showed signs of a blunted endothelial response similar to that of their infants. Accordingly, contributions from a common genetic pathway for endothelial dysfunction in LBW infants cannot be excluded.

Besides LBW, this study indicates low folate levels as a risk factor for endothelial dysfunction. Preterm infants do not show impaired endothelial function.27 This may indicate that the third trimester is a particularly sensitive period for adverse vascular effects. Therefore, adequate folate levels may be important throughout pregnancy until term and not only during the periconceptional period for prevention of neural tube defects.

The possibility of deficient placental transfer of folate as a cause of fetal vascular damage is contradicted by our findings. In agreement with previous data on a concentration gradient across the placenta,20 we and others previously found that folate levels in the offspring exceeded those in the mother.17,18,29

All infants in our study were breastfed. To prevent hypoglycemia, LBW infants more often received extra formula in the first days after birth. Therefore, the possibility that LBW infants had higher postnatal folate intake cannot be excluded. However, because mothers of LBW infants showed lower folate levels, their breast milk likely contained less folate. In addition, RBC folate levels are not influenced by daily fluctuations in intake but reflect more-long-term bioavailability.

For almost all mothers and infants in our study, the folate levels were within the reference range defined by our chemistry laboratory (RBC folate levels of 350–1500 nmol/L). This could mean that not only folate-deficient pregnant women are at risk for adverse vascular effects on their fetuses. In fact, we found that the differences in neonatal endothelial function were related to physiologic variations in maternal folate levels, which can be much larger than those reported here.30

The molecular mechanisms through which low folate levels could acutely affect endothelial function include oxidative inactivation and reduced synthesis of the endothelium-derived vasodilator nitric oxide (Fig 1).12,13 These effects are reversible once folate levels are increased. Because endothelial function was measured only a few days after birth, maternal folate levels could still have been relevant to infant vascular function. Low folate levels have also been found to relate to hyperhomocysteinemia, which may add to endothelial dysfunction.12,13 However, our results do not indicate that the association between folate levels and endothelial function in newborn infants is mediated via high homocysteine levels.

Although previous cross-sectional studies indicated that endothelial dysfunction in LBW subjects persists throughout childhood and beyond,6–10 it is possible that the presently found association between low folate levels and endothelial dysfunction may be transient. Questions regarding if and how folate levels in early life can have long-lasting effects on the vascular system remain to be clarified. There is a need for longitudinal studies beyond the neonatal period to determine the relevance of our findings to cardiovascular disease in adulthood. In addition, the role of folate in DNA methylation, with genomic imprinting and epigenetic activation or silencing of genes, could be one target for future mechanistic research.

In a rat model, endothelial nitric oxide synthase mRNA levels were found to be decreased in adult offspring of protein-restricted dams. Folate supplementation for the protein-restricted pregnant rats increased endothelial nitric oxide synthase expression in the adult offspring (Fig 1).26

In Sweden, grain products are not enriched with folic acid. After folic acid fortification of grains was implemented in the United States and Canada in 1998, stroke mortality rates have been observed to decrease.31 Although secondary prevention with folic acid and vitamin B12 supplementation for middle-aged cardiovascular patients failed to lower the risk of recurrent disease,12,33 the impact of folate supplementation during pregnancy may be different.

In previous studies of LBW infants and children, we demonstrated that vasodilation independent of the endothelium was unaffected.6,8 Given the short periods of quiet sleep for newborn infants, which offer a limited opportunity for vascular measurements, only tests of endothelial function were performed in this study.

CONCLUSIONS

Our findings suggest a relationship between low maternal folate levels in late pregnancy and vascular endothelial dysfunction in the newborn infants, independent of LBW. Low maternal folate levels may be a preventable perinatal contribution to cardiovascular disease in the offspring. Studies of maternal, placental, and fetal outcomes after folate and vitamin B12 supplementation throughout pregnancy are warranted. Longitudinal follow-up studies of subjects born to mothers participating in previous or coming trials of macronutrient and micronutrient interventions during pregnancy would also be helpful. Taken together, such efforts should contribute to useful, evidence-based, perinatal strategies for prevention of adult disease and improvements in public health.

ACKNOWLEDGMENTS

This study was supported by grants from the Swedish Research Council (project 71P-14158 and project 348-2002-6975), Swedish Heart Lung Foundation, Karolinska Institute Research Foundations, and Sällskapet Barnavård, Karolinska University Hospital.

We thank Jessica Schiött, research nurse, for help with the study.
REFERENCES


Inhaled Nitric Oxide Therapy Decreases the Risk of Cerebral Palsy in Preterm Infants With Persistent Pulmonary Hypertension of the Newborn

Yuko Tanaka, MD, Tomoshige Hayashi, MD, PhD, Hiroyuki Kitajima, MD, PhD, Kiyaki Sumi, MD, Masanori Fujimura, MD, PhD

Department of Neonatal Medicine, Osaka Medical Center and Research Institute for Maternal and Child Health, Osaka, Japan; Department of Preventive Medicine and Environmental Health, Graduate School of Medicine, Osaka City University, Osaka, Japan

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

ABSTRACT

OBJECTIVE. The aim was to determine whether inhaled nitric oxide therapy improves neurodevelopmental outcomes for infants with preterm persistent pulmonary hypertension of the newborn.

METHODS. We conducted a historical cohort study to compare the 3-year incidence of cerebral palsy in preterm singleton infants (<34 gestational weeks) with hypoxic respiratory failure caused by persistent pulmonary hypertension of the newborn who received inhaled nitric oxide (16 patients) or 100% oxygen (15 patients) therapy. All neonates had clinical and echocardiographic evidence of pulmonary hypertension without structural heart disease.

RESULTS. The incidence of cerebral palsy among patients treated with inhaled nitric oxide was 12.5%, whereas that among patients treated with 100% oxygen was 46.7%. After adjustment for maternal fever (≥38°C) during delivery, birth weight, Apgar score at 5 minutes, high-frequency oscillatory ventilation, and surfactant therapy, inhaled nitric oxide therapy, compared with 100% oxygen therapy, was associated with a decreased risk of cerebral palsy in preterm infants with persistent pulmonary hypertension of the newborn.

CONCLUSION. Inhaled nitric oxide therapy decreases the risk of cerebral palsy in preterm infants with persistent pulmonary hypertension of the newborn.
Persistent Pulmonary Hypertension of the newborn (PPHN) is a disease in which the pulmonary vascular resistance remains elevated during the neonatal period. Preterm PPHN, which is associated with a high risk of adverse health and neurodevelopmental outcomes, continues to be one of the most challenging conditions encountered in the NICU.

Nitric oxide (NO) is produced by vascular endothelial cells and plays an important role in increasing the blood flow to the lungs after birth.\(^1\)\(^-\)\(^5\) Inhaled NO (iNO) causes selective pulmonary vasodilation in newborn lambs\(^1\) and has been shown to have a short-term benefit by improving oxygenation in infants with hypoxemic respiratory failure caused by PPHN.\(^6\) However, it is not clear whether, in preterm infants with PPHN, iNO therapy decreases the risk of cerebral palsy (CP), which is one of the most serious neurodevelopmental complications of preterm infants. We conducted a historical cohort study to compare the 3-year incidence of CP in preterm singleton infants with hypoxemic respiratory failure caused by PPHN who received either iNO or 100% oxygen therapy.

**METHODS**

**Study Subjects**

A historical cohort study was performed that involved 61 consecutive patients without congenital anomalies who were admitted to the ICU at the Osaka Medical Center and Research Institute for Maternal and Child Health between January 1988 and December 1999 and who were singleton infants of <34 gestational weeks (median: 25.0 weeks; interquartile range [IQR]: 24.0–28.0 weeks) with hypoxemic respiratory failure caused by PPHN. Thirty of the 61 patients in the original cohort were excluded; 26 patients died within 3 years after birth (median: 1.0 day; IQR: 0.6–7.8 days) and 4 patients were weaned off iNO therapy. Infants who exhibited improvement continued to receive iNO at the minimal level found to be effective (attempts were made to decrease the concentration by reducing it by 5 ppm at 30-minute intervals). At that time, the fraction of inspired oxygen (Fio\(_2\)) and ventilation were reduced to prevent additional lung injury. When the Fio\(_2\) could be decreased to ≤0.4, iNO therapy was terminated by slow weaning over several hours. Methemoglobin levels were measured before iNO therapy, 1 hour later, and then every 8 hours. If the methemoglobin level increased to >2%, then iNO therapy was discontinued. Infants were monitored for signs of increased bleeding (eg, pulmonary hemorrhage, gastrointestinal bleeding, or oozing from venipuncture sites). In addition, cranial ultrasonography was performed before the initiation, within 24 hours whenever possible and then every 24 hours after the initiation, and 24 hours after the final discontinuation of iNO therapy. The median duration of iNO therapy was 19.8 hours (IQR: 29.5–56.0 hours). The oxygenation index, calculated as \([100 \times Fio_2 \times mean\ airway\ pressure\ (in\ centimeters\ of\ water)]/PaO_2\) (in millimeters of mercury), was obtained within 1 hour before and at 1 hour after the initiation of inhalation therapy (iNO or 100% oxygen). The mothers’ records were re-

The study protocol was in accordance with the institutional guidelines for human research, and the patients’ parents provided written informed consent for the diagnostic and therapeutic procedures that were required, which allowed the results of the examinations to be used in this study.

**Data Collection**

Dr Tanaka reviewed the infant and maternal records. Between January 1988 and September 1993, all preterm infants with PPHN received 100% oxygen therapy to treat their respiratory failure. In 1992, iNO therapy was reported to improve respiratory failure in patients with PPHN;\(^7\) therefore, after institutional ethics committee approval was obtained, iNO therapy was given to all preterm infants with PPHN between October 1993 and December 1999. During the 2 time periods, the patients received similar treatments except for the iNO therapy or 100% oxygen therapy.

NO gas (Taiyo Toyo Sanso, Osaka, Japan) was delivered from an 800-ppm cylinder and was introduced into the afferent limb of the ventilator circuit near the endotracheal tube, which mixed the fixed flow of gas in the ventilator circuit. The flow was adjusted to yield the predetermined NO concentration. The iNO concentration was increased by 10 ppm at 30-minute intervals, with an upper limit of 30 ppm. The response to iNO therapy was evaluated as an increase in Pao\(_2\) to >10 mm Hg. Infants who did not show a significant acute response continued to receive iNO therapy at 5 ppm for 12 hours; if there was still no satisfactory response, then the infants were weaned off iNO therapy. Infants who exhibited improvement continued to receive iNO at the minimal level found to be effective (attempts were made to decrease the concentration by reducing it by 5 ppm at 30-minute intervals).

**RESULTS**

**PPHN**

Persistently elevated pulmonary artery (PA) pressure, which exceeds the systemic arterial pressure, is considered to be the major cause of right-to-left shunting. However, the presence of a significant right-to-left shunt was considered to be a contraindication for iNO therapy.

**Umbilical Artery Oxygen Saturation**

Umbilical artery saturation is the most commonly used measure of oxygenation in the neonatal period. It is best determined by transcutaneous monitoring during the first hour of life. However, the umbilical artery sample is not always available.

**Cyanosis**

Cyanosis is a common finding in infants with PPHN. It is usually caused by an increase in the fraction of inspired oxygen (Fio\(_2\)) and ventilation. However, cyanosis can also be caused by a decrease in the fraction of inspired oxygen (Fio\(_2\)) and ventilation.

**PPHN Grading System**

No standard grading system is currently available. However, a grading system based on the severity of the disease is often used. This grading system is based on the degree of hypoxemia, the presence of right-to-left shunting, and the response to treatment.

**Monitoring**

Monitoring is an important aspect of the management of PPHN. It is important to monitor the patient’s oxygenation, heart rate, and blood pressure. Monitoring of arterial blood gas values and oxygen saturation is also important.

**Conclusion**

In conclusion, iNO therapy is a feasible and effective treatment for PPHN. However, it is important to monitor the patient’s oxygenation, heart rate, and blood pressure. Monitoring of arterial blood gas values and oxygen saturation is also important.

**Acknowledgments**

This study was supported by grants from the Ministry of Health, Labor, and Welfare of Japan and the Ministry of Education, Culture, Sports, Science, and Technology of Japan.
viewed to determine the presence of maternal fever (≥38°C) during delivery, premature rupture of the membranes (≥24 hours), maternal bleeding, reason for delivery, and prenatal corticosteroid use. Neonatal data, which were obtained from the medical charts, included the number of gestational weeks at birth, gender, birth weight, Apgar scores at 1 and 5 minutes, use of iNO, type of ventilation (high-frequency oscillatory or intermittent mechanical ventilation), and surfactant therapy. All of the subjects’ parents were Japanese. All surviving infants were scheduled to be seen by pediatricians at 3 years of age for a complete physical and neurologic examination. Necrotizing enterocolitis was diagnosed during surgery, at autopsy, or on the basis of a finding of pneumatosi

Statistical Analyses

Categorical variables were compared by using the χ² test or Fisher’s exact test. Differences in the median values between the 2 groups were compared by using the Mann-Whitney U test. Univariate and multivariate logistic regression analyses were used to estimate the odds ratio for the incidence of CP. We calculated the 95% confidence interval for each odds ratio. We limited the number of independent variables in each model to avoid overfitting the data. P values were 2-tailed. A P value of <.05 was considered significant. Statistical analyses were performed by using the SPSS 10.0 software package (SPSS, Chicago, IL).

RESULTS

Of the 61 preterm infants with PPHN, 26 infants died within 3 years after birth. Mortality rates at 3 years after birth were similar for infants treated with iNO and those treated with 100% oxygen (44.1% vs 40.7%; P = .791). The incidences of outcomes in the iNO-treated and 100% oxygen-treated groups were 8.8% vs 29.6% for patent ductus arteriosus, 17.6% vs 25.9% for intraventricular hemorrhage (grade 3 or 4), 8.8% vs 0% for necrotizing enterocolitis, and 11.8% vs 16.0% for pulmonary hemorrhage.

During the 3-year follow-up period, the incidence of CP among patients treated with iNO therapy was lower than that among patients treated with 100% oxygen therapy (12.5% vs 46.7%; P = .054). The baseline clinical characteristics of the study patients according to type of inhalation therapy are summarized in Table 1. Patients who received iNO therapy had lower Apgar scores at 5 minutes and were given surfactant therapy more often, compared with those who received 100% oxygen therapy (Table 1). In univariate logistic analysis, use of iNO therapy was associated with a decreased incidence of CP (Table 2).

We tested several regression models to assess the effect of iNO therapy on the incidence of CP in preterm infants with PPHN. After adjustments for maternal fever during delivery, birth weight, Apgar score at 5 minutes, high-frequency oscillatory ventilation, and surfactant therapy, iNO therapy was associated with a decreased risk of CP, compared with 100% oxygen therapy (Table 3).

DISCUSSION

Our data demonstrate that, for preterm infants with PPHN, the incidence of CP, during a 3-year follow-up

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Baseline Characteristics of Patients According to Treatment Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>iNO (N = 16)</td>
</tr>
<tr>
<td><strong>Maternal characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>Maternal fever (≥38°C) during delivery, n (%)</td>
<td>2 (13.3)</td>
</tr>
<tr>
<td>Premature rupture of membranes (≥24 h), n (%)</td>
<td>4 (26.7)</td>
</tr>
<tr>
<td>Maternal bleeding as reason for delivery, n (%)</td>
<td>10 (66.7)</td>
</tr>
<tr>
<td>Prenatal corticosteroid therapy, n (%)</td>
<td>2 (13.3)</td>
</tr>
<tr>
<td><strong>Infant characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>Gestational age, median (IQR), wk</td>
<td>25.5 (25.0–28.8)</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>8 (50.0)</td>
</tr>
<tr>
<td>Birth weight, median (IQR), g</td>
<td>838 (628–1144)</td>
</tr>
<tr>
<td>Apgar score at 1 min, median (IQR)</td>
<td>4.0 (1.5–4.8)</td>
</tr>
<tr>
<td>Apgar score at 5 min, median (IQR)</td>
<td>6.0 (5.0–7.0)</td>
</tr>
<tr>
<td>High-frequency oscillatory ventilation, n (%)</td>
<td>14 (87.5)</td>
</tr>
<tr>
<td>Surfactant used, n (%)</td>
<td>15 (93.8)</td>
</tr>
</tbody>
</table>

* Data are missing for 1 infant.
period, among patients treated with iNO therapy showed a trend toward being lower than that among patients treated with 100% oxygen therapy. After adjustments for maternal fever during delivery, birth weight, Apgar score at 5 minutes, high-frequency oscillatory ventilation, and surfactant therapy, iNO therapy was associated with a decreased risk of CP, compared with 100% oxygen therapy.

Previous research had not shown that iNO therapy reduces the risk of CP in preterm infants with hypoxic respiratory failure. Bennett et al\(^8\) reported that, in a randomized, controlled study of 42 preterm infants who were thought to be at high risk of developing chronic lung disease, the rates of CP at 30 months of age were similar in the iNO-treated and control groups. In addition, Mestan et al\(^9\) conducted a double-blind, randomized, controlled trial of 138 preterm infants with respiratory failure and found that, although iNO therapy improved cognitive neurodevelopmental outcomes, the rates of CP at 2 years of age were similar in the iNO-treated and control groups. Those results are not consistent with our findings. The study groups in the previous studies were not limited to preterm infants with PPHN, whereas we included only preterm infants with PPHN in the present study. Therefore, the reduced risk of CP with iNO therapy may be limited to such cases. However, because the incidence of CP in the control group was low in the previous studies (14% in the study by Bennett et al\(^8\) and 10% in the study by Mestan et al\(^9\)), compared with our study (47%), the beneficial neurodevelopmen-

### TABLE 2  Baseline Characteristics of Patients According to Whether CP Developed After 3 Years

<table>
<thead>
<tr>
<th>Variable</th>
<th>Yes (n = 9)</th>
<th>No (n = 22)</th>
<th>Crude Odds Ratio (95% Confidence Interval)</th>
<th>(P^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maternal fever ((\geq 38^\circ)) during delivery, n (%)</td>
<td>3 (37.5)(^a)</td>
<td>5 (22.7)</td>
<td>2.04 (0.36–11.67)</td>
<td>.423</td>
</tr>
<tr>
<td>Premature rupture of membranes ((\geq 24) h), n (%)</td>
<td>2 (22.2)</td>
<td>5 (23.8)(^b)</td>
<td>0.91 (0.14–5.90)</td>
<td>.925</td>
</tr>
<tr>
<td>Maternal bleeding as reason for delivery, n (%)</td>
<td>5 (62.5)(^b)</td>
<td>14 (63.6)</td>
<td>0.95 (0.18–5.08)</td>
<td>.954</td>
</tr>
<tr>
<td>Prenatal corticosteroid therapy, n (%)</td>
<td>3 (37.5)(^b)</td>
<td>4 (18.2)</td>
<td>2.70 (0.45–16.26)</td>
<td>.278</td>
</tr>
<tr>
<td>Infant characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gestational age, median (IQR), wk</td>
<td>26.0 (25.0–28.0)</td>
<td>26.0 (24.0–29.8)</td>
<td>0.97 (0.74–1.29)(^c)</td>
<td>.848</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>5 (55.6)</td>
<td>11 (50.0)</td>
<td>1.25 (0.26–5.94)</td>
<td>.779</td>
</tr>
<tr>
<td>Birth weight, median (IQR), g</td>
<td>796 (752–880)</td>
<td>914 (638–1303)</td>
<td>0.52 (0.15–1.81)(^d)</td>
<td>.307</td>
</tr>
<tr>
<td>Apgar score at 1 min, median (IQR)</td>
<td>3.0 (1.0–7.0)</td>
<td>4.0 (3.0–5.0)</td>
<td>0.99 (0.66–1.50)</td>
<td>.978</td>
</tr>
<tr>
<td>Apgar score at 5 min, median (IQR)</td>
<td>7.0 (4.0–9.0)</td>
<td>7.0 (6.0–7.0)</td>
<td>0.82 (0.55–1.23)</td>
<td>.336</td>
</tr>
<tr>
<td>iNO used, n (%)</td>
<td>2 (22.2)</td>
<td>14 (63.6)</td>
<td>0.16 (0.03–0.98)</td>
<td>.048</td>
</tr>
<tr>
<td>High-frequency oscillatory ventilation, n (%)</td>
<td>8 (89)</td>
<td>15 (68)</td>
<td>3.73 (0.39–35.84)</td>
<td>.254</td>
</tr>
<tr>
<td>Surfactant used, n (%)</td>
<td>7 (77.8)</td>
<td>17 (77.3)</td>
<td>1.03 (0.16–6.62)</td>
<td>.976</td>
</tr>
</tbody>
</table>

\(a\) \(P\) values for univariate logistic analyses.

\(b\) Data are missing for 1 infant.

\(c\) Per 1-week increase.

\(d\) Per 500-g increase.

### TABLE 3  Multivariate Models for Determining the Incidence of CP

<table>
<thead>
<tr>
<th>Model</th>
<th>Variable</th>
<th>Odds Ratio (95% Confidence Interval)</th>
<th>(P^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>iNO therapy vs 100% oxygen therapy</td>
<td>0.08 (0.01–0.85)</td>
<td>.036</td>
</tr>
<tr>
<td></td>
<td>Maternal fever ((\geq 38^\circ)) during delivery</td>
<td>1.07 (0.15–7.52)</td>
<td>.947</td>
</tr>
<tr>
<td>Model 2</td>
<td>iNO therapy vs 100% oxygen therapy</td>
<td>0.12 (0.02–0.84)</td>
<td>.032</td>
</tr>
<tr>
<td></td>
<td>Birth weight, per 500-g increase</td>
<td>0.38 (0.09–1.57)</td>
<td>.181</td>
</tr>
<tr>
<td>Model 3</td>
<td>iNO therapy vs 100% oxygen therapy</td>
<td>0.06 (0.05–0.70)</td>
<td>.025</td>
</tr>
<tr>
<td></td>
<td>Apgar score at 5 min</td>
<td>0.59 (0.32–1.09)</td>
<td>.092</td>
</tr>
<tr>
<td>Model 4</td>
<td>iNO therapy vs 100% oxygen therapy</td>
<td>0.08 (0.01–0.61)</td>
<td>.014</td>
</tr>
<tr>
<td></td>
<td>High-frequency oscillatory ventilation vs intermittent mechanical ventilation</td>
<td>10.5 (0.85–130.51)</td>
<td>.067</td>
</tr>
<tr>
<td>Model 5</td>
<td>iNO therapy vs 100% oxygen therapy</td>
<td>0.12 (0.02–0.82)</td>
<td>.031</td>
</tr>
<tr>
<td></td>
<td>Surfactant therapy</td>
<td>2.67 (0.33–21.68)</td>
<td>.359</td>
</tr>
</tbody>
</table>

\(a\) \(P\) values for multivariate logistic analyses.
tal outcomes associated with iNO treatment might have been underestimated in the previous studies; the incidence of CP in the iNO-treated group was 0% in the study by Bennett et al and 9% in the study by Mestan et al.9

We did not identify the reasons why iNO therapy decreased the risk of CP in preterm infants with PPHN. However, an in vitro study using rat brain slices showed that hypoxia induced white matter damage mainly through oxidation.10 In comparison with gray matter, white matter contains larger amounts of fat and iron, which are involved in free radical production, and a smaller amount of glutathione, which is an antioxidant; this suggests that white matter has greater susceptibility to oxidative stress. In addition, blood flow to cerebral white matter is extremely low in premature newborns,1,1 which indicates that cerebral white matter is particularly vulnerable to hypoxia in preterm infants. These results suggest that hypoxia easily can induce white matter damage in preterm infants. We demonstrated that, although the oxygenation index values before the start of inhalation therapy were similar in the iNO-treated and 100% oxygen-treated groups, the oxygenation index 1 hour after the start of inhalation therapy was lower for preterm infants with PPHN treated with iNO, compared with infants treated with 100% oxygen. Therefore, iNO therapy may decrease the risk of CP in these infants through the resolution of hypoxia during a critical phase of neurodevelopment. Alternatively, iNO therapy may affect the brain directly by stimulating neuronal maturation.12-14 However, there is no clear evidence that iNO affects brain development directly.

Our study has some potential limitations. First, we did not conduct a randomized, placebo-controlled trial, for ethical reasons; we conducted a historical cohort study to compare the 3-year incidence of CP in preterm infants with PPHN who received either iNO or 100% oxygen therapy. After adjustment for multiple potential confounding variables, iNO therapy was associated with a decreased risk of CP. In addition, the type of inhalation therapy was determined on the basis of the time of each subject’s admission. Therefore, the selection of the type of inhalation can not introduce bias. Second, in our study, we enrolled only preterm infants with PPHN. Therefore, it is not clear whether iNO therapy, compared with 100% oxygen therapy, would decrease the risk of CP in preterm infants with hypoxic respiratory failure not caused by PPHN. Third, we analyzed a limited number of patients. Therefore, other variables, such as early gestational age, low birth weight, presence of maternal fever during delivery,15,16 premature rupture of membranes of long duration,16-20 maternal bleeding, reason for preterm delivery,21,22 and low Apgar scores at birth,16,17,22 which are thought be risk factors for CP, might not have been identified as being useful for predicting CP in this study. In addition, use of surfactant therapy and use of high-frequency oscillatory ventilation were not associated with a decreased incidence of CP in preterm infants with PPHN. These results might be attributable to a limited number of study subjects. In particular, the use of high-frequency oscillatory ventilation had a high odds ratio, compared with intermittent mechanical ventilation. Therefore, to generalize the results of this study, studies involving a large number of patients are essential. Finally, because we conducted a historical cohort study, we could not explore unknown risk factors for CP.

CONCLUSIONS

Our results provide evidence that iNO therapy, compared with 100% oxygen therapy, decreases the risk for CP in preterm infants with PPHN. iNO therapy may protect brain white matter during a critical phase of neurodevelopment and thus reduce the risk of CP in these infants. Additional studies are needed to clarify the mechanism through which iNO therapy decreases the risk of CP in preterm infants with PPHN.

ACKNOWLEDGMENTS

We thank the staff members, patients, and parents who took part in this study. We thank Dr Luba Wolchuk for correcting our manuscript.

REFERENCES

5. Steinhorn RH, Morin FC III, Fineman JR. Models of persistent pulmonary hypertension of the newborn (PPHN) and the role of cyclic guanosine monophosphate (GMP) in pulmonary vasorelaxation. Semin Perinatol. 1997;21:393–408


17. Nelson KB, Ellenberg JH. Predictors of low and very low birth weight and the relation of these to cerebral palsy. *JAMA.* 1985;254:1473–1479


---

**DROTRECOGIN ALFA (ACTIVATED) IN CHILDREN WITH SEVERE SEPSIS: A MULTICENTRE PHASE III RANDOMISED CONTROLLED TRIAL**

**Background:** Drotrecogin alfa (activated) (DrotAA) is used for the treatment of adults with severe sepsis who have a high risk of dying. A phase 1b open-label study has indicated that the pharmacokinetics and pharmacodynamics of DrotAA are similar in children and adults. We initiated the RESOLVE (REsearching severe Sepsis and Organ dysfunction in children: a gLocal perspectiVE) trial to investigate the efficacy and safety of the drug in children.

**Methods:** Children aged between 38 weeks’ corrected gestational age and 17 years with sepsis-induced cardiovascular and respiratory failure were randomly assigned to receive placebo or DrotAA (24 μg/kg/h) for 96 h.

**Findings:** 477 patients were enrolled: 237 received placebo, and 240 DrotAA. Our results showed no significant difference between groups.

**Interpretation:** Although we did not record any efficacy of DrotAA in children with severe sepsis, serious bleeding events were similar between groups and the overall safety profile acceptable, except in children younger than 60 days. However, we gained important insights into clinical and laboratory characteristics of childhood severe sepsis, and have identified issues that need to be addressed in future trials in critically ill children.


Noted by JFL, MD
Patent Ductus Arteriosus and Its Treatment as Risk Factors for Neonatal and Neurodevelopmental Morbidity

Nancy Chorne, MDa, Carol Leonard, PhDa, Robert Piecuch, MDa, Ronald I. Clyman, MDa,b

aDepartment of Pediatrics and bCardiovascular Research Institute, University of California, San Francisco, California

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

ABSTRACT

OBJECTIVES. The purpose of this work was to determine whether the reported association between neonatal morbidities and a patent ductus arteriosus is because of the left-to-right patent ductus arteriosus shunt itself, the therapies used to treat it, or the immaturity of the infants who are likely to develop a patent ductus arteriosus.

METHODS. A total of 446 infants (<28 weeks’ gestation) were treated with the same patent ductus arteriosus care–oriented protocol, and logistic regression analysis was used to examine the effects of several patent ductus arteriosus–related variables (presence of a symptomatic patent ductus arteriosus, the number of indomethacin doses used, the ductus response to indomethacin, and the use of surgical ligation) on the incidence of retinopathy of prematurity, necrotizing enterocolitis, chronic lung disease, death, and neurodevelopmental impairment.

RESULTS. Most of the predictive effects that the presence of a patent ductus arteriosus and its treatment had on neonatal morbidity could be accounted for by the infants’ immature gestation. Use of surgical ligation, however, was significantly associated with the development of chronic lung disease and was independent of immature gestation, other patent ductus arteriosus–related variables, or other perinatal and neonatal risk factors known to be associated with chronic lung disease.

CONCLUSIONS. These findings add to the growing uncertainty about the benefits and risks of surgical ligation during the neonatal period.
The presence of a persistent left-to-right shunt through a patent ductus arteriosus (PDA) is associated with the development of other neonatal morbidities. At this time it is unclear whether the reported association between a persistent PDA and other neonatal morbidities is because of the left-to-right PDA shunt itself, the therapies used to treat it, or the immaturity of the infant who is likely to develop it. Most studies examining this issue have looked only at individual aspects of PDA-oriented care on neonatal morbidity (eg, the presence of a PDA by itself [without regard for size of shunt or treatments used], duration of exposure to a symptomatic PDA, use of indomethacin, timing of indomethacin treatment, duration of indomethacin treatment, ductus response to indomethacin treatment, and use of surgical ligation). Few have considered all of the variables together when determining the effect of a PDA or its therapies on neonatal morbidity.

In the following study, we used a cohort of infants (<28 weeks' gestation) who were managed with a defined PDA care-oriented protocol in a single medical center to examine the effects of several aspects of PDA-oriented care on neonatal morbidity. We used multiple logistic regression modeling to determine which aspects of PDA-oriented care were most closely associated with the development of other neonatal morbidities.

**METHODS**

**Population and PDA Treatment Protocol**

This project was approved by the institutional review board of the University of California San Francisco. Between January 1994 and July 2005, all of the infants <28 weeks' gestation, admitted within 15 hours of birth to the William H. Tooley Nursery at University of California San Francisco, were treated according to the following PDA care-oriented protocol. Infants received either a short 3-dose course of prophylactic indomethacin (0.2, 0.1, and 0.1 mg/kg, administered at 24-hour intervals) or an extended 6-dose course (0.2, 0.1, 0.1, 0.1, 0.1, and 0.1 mg/kg, at 24-hour intervals) starting within 15 hours of birth. A Doppler examination was performed after the second dose. If there was no evidence of ductus patency on the Doppler examination, infants received the short 3-dose course (prophylactic indomethacin was stopped after the third dose). If there was any evidence of ductus patency on the examination, infants received the extended 6-dose course. A repeat echo-Doppler examination was performed 24 to 36 hours after the last dose to determine the ductus response to prophylactic indomethacin.

After the prophylactic treatment, infants were examined daily for the appearance of clinical symptoms indicative of a PDA (systolic murmur, widened pulse pressure, and hyperdynamic precordium). If any of these occurred, an echo-Doppler examination was performed within 24 hours. If there was left-to-right flow through the PDA, the infant was considered to have a symptomatic PDA. The degree of left-to-right shunt was considered to be small or moderate on the basis of the absence or presence of holodiastolic retrograde flow in the descending aorta (at the level of the diaphragm). When infants developed a symptomatic PDA, they were treated with a 3-dose course of indomethacin (0.2, 0.1, and 0.1 mg/kg administered at 0, 12, and 36 hours) and/or ligation. The need for treatment did not depend on the need for respiratory support or the degree of left-to-right shunt. Even infants who required only nasal cannula oxygen received treatment if the ductus was patent on echo-Doppler examination.

The choice of which treatment (indomethacin or surgery) to initiate first for the symptomatic PDA was left to the attending neonatologist. The decision to use surgery, rather than indomethacin, was based mainly on the ductus response to the initial prophylactic indomethacin course, not on the infant's medical condition. Infants who closed their PDA after prophylactic indomethacin (by echo/Doppler) were more likely to be treated with a second 3-dose course of indomethacin if a symptomatic PDA developed; those who still had evidence of Doppler flow after the course of prophylactic indomethacin, were more likely to be sent directly to surgery if the PDA became symptomatic. The infant's medical condition sometimes played a role in the treatment choice. Indomethacin was more likely to be used, as a first-line therapy, if infants were unstable when their symptomatic PDA presented; surgery was preferred for more stable infants. Eighteen percent of symptomatic PDAs were treated with indomethacin alone, 13% with ligation alone, and 69% with indomethacin and ligation. The age of presentation of symptomatic PDAs that were treated with indomethacin (15 ± 7 days [mean ± SD]) was not different from those that were treated with surgery (14 ± 8 days). All of the symptomatic PDAs were closed, pharmacologically and/or surgically, within 5 days of presentation. The duration of exposure to a symptomatic PDA, the interval between symptomatic PDA presentation and the start of indomethacin treatment, was 0.8 ± 1.3 days (mean ± SD; an echo to document closure was performed 2 days later). The interval between symptomatic PDA presentation and ligation was 1.9 ± 1.2 days (mean ± SD). This approach to PDA-oriented care was based on a review of the randomized, controlled trials that examined the relative effectiveness and morbidities of indomethacin prophylaxis and of prolonged exposure to a symptomatic PDA.

**Risk Factors and Outcome Variables**

The perinatal and neonatal characteristics of this population are listed in Table 1. Gestational age was determined by the date of last menstrual period and early
### TABLE 1  Demographics

<table>
<thead>
<tr>
<th>Variables</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Perinatal/neonatal variables (n = 446)</strong></td>
<td></td>
</tr>
<tr>
<td>Gestation, wk</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>0.9</td>
</tr>
<tr>
<td>24</td>
<td>19.7</td>
</tr>
<tr>
<td>25</td>
<td>24.9</td>
</tr>
<tr>
<td>26</td>
<td>24.7</td>
</tr>
<tr>
<td>27</td>
<td>29.8</td>
</tr>
<tr>
<td>Year of birth 2000 or later&lt;sup&gt;a&lt;/sup&gt;</td>
<td>53.1</td>
</tr>
<tr>
<td>Male gender</td>
<td>52.8</td>
</tr>
<tr>
<td>Betamethasone exposure</td>
<td></td>
</tr>
<tr>
<td>None or 6h</td>
<td>26.8</td>
</tr>
<tr>
<td>6–23 h</td>
<td>14.9</td>
</tr>
<tr>
<td>≥24 h</td>
<td>58.3</td>
</tr>
<tr>
<td>Preeclampsia</td>
<td>22.3</td>
</tr>
<tr>
<td>Gestational diabetes</td>
<td>6.0</td>
</tr>
<tr>
<td>Chorioamnionitis</td>
<td>21.6</td>
</tr>
<tr>
<td>Respiratory distress</td>
<td>92.0</td>
</tr>
<tr>
<td>Surfactant</td>
<td>94.6</td>
</tr>
<tr>
<td>Respiratory score &gt;3&lt;sup&gt;b&lt;/sup&gt;</td>
<td>16.7</td>
</tr>
<tr>
<td>ICH</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>44.9</td>
</tr>
<tr>
<td>Grade 1</td>
<td>25.5</td>
</tr>
<tr>
<td>Grade 2</td>
<td>15.2</td>
</tr>
<tr>
<td>Grade 3</td>
<td>4.0</td>
</tr>
<tr>
<td>Grade 4</td>
<td>10.5</td>
</tr>
<tr>
<td>Sepsis&lt;sup&gt;c&lt;/sup&gt;</td>
<td>42.5</td>
</tr>
<tr>
<td>Average daily fluids &gt;160 mL/kg per day&lt;sup&gt;e&lt;/sup&gt;</td>
<td>19.6</td>
</tr>
<tr>
<td>NEC&lt;sup&gt;d&lt;/sup&gt;</td>
<td>15.8</td>
</tr>
<tr>
<td>Periventricular leukomalacia&lt;sup&gt;d&lt;/sup&gt;</td>
<td>4.7</td>
</tr>
<tr>
<td>CLD&lt;sup&gt;d&lt;/sup&gt;</td>
<td>27.5</td>
</tr>
<tr>
<td>ROP&lt;sup&lt;d&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Stage 2 with plus disease or ≥ stage 3</td>
<td>15.0</td>
</tr>
<tr>
<td>Laser treatment</td>
<td>10.9</td>
</tr>
<tr>
<td>Death</td>
<td></td>
</tr>
<tr>
<td>During initial hospitalization</td>
<td>19.1</td>
</tr>
<tr>
<td>After discharge</td>
<td>2.2</td>
</tr>
<tr>
<td>Maternal education (&lt;12 years completed)&lt;sup&gt;f&lt;/sup&gt;</td>
<td>26.3</td>
</tr>
<tr>
<td><strong>PDA-oriented variables (n = 446)</strong></td>
<td></td>
</tr>
<tr>
<td>Indomethacin exposure (number of indomethacin doses)&lt;sup&gt;g&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Prophylactic treatment doses ≤3</td>
<td>72.8</td>
</tr>
<tr>
<td>Prophylactic treatment doses &gt;3</td>
<td>27.2</td>
</tr>
<tr>
<td>Total number during hospitalization ≤3</td>
<td>66.5</td>
</tr>
<tr>
<td>Total number during hospitalization &gt;3</td>
<td>33.5</td>
</tr>
<tr>
<td>Ductus patent after prophylactic indomethacin&lt;sup&gt;h&lt;/sup&gt;</td>
<td>14.7</td>
</tr>
<tr>
<td>Symptomatic PDA&lt;sup&gt;i&lt;/sup&gt;</td>
<td>26.5</td>
</tr>
<tr>
<td>Ligation&lt;sup&gt;d&lt;/sup&gt;</td>
<td>22.4</td>
</tr>
<tr>
<td><strong>Neurodevelopmental outcome (n = 369)</strong></td>
<td></td>
</tr>
<tr>
<td>Death or neurodevelopmental impairment&lt;sup&gt;j&lt;/sup&gt;</td>
<td>36.6</td>
</tr>
<tr>
<td>Cerebral palsy&lt;sup&gt;k&lt;/sup&gt;</td>
<td>5.2</td>
</tr>
<tr>
<td>Neurosensory (vision/hearing) impairment&lt;sup&gt;k&lt;/sup&gt;</td>
<td>6.7</td>
</tr>
<tr>
<td>Cognitive delay&lt;sup&gt;k&lt;/sup&gt;</td>
<td>12.0</td>
</tr>
<tr>
<td>Cognitive delay (15–30 mo)&lt;sup&gt;l&lt;/sup&gt;</td>
<td>18.0</td>
</tr>
</tbody>
</table>

---

<sup>a</sup> Birth date was on or after January 1, 2000.
<sup>b</sup> Respiratory score (mean airway pressure × fraction of inspired oxygen) was at 24 hours after birth (population mean: 2.25; range: 0.84–14.00).
<sup>c</sup> Sepsis indicates symptoms and positive blood, urine, or cerebrospinal fluid culture.
<sup>d</sup> Infants with early death were not included in the calculation because outcome could not be determined.
<sup>e</sup> Average daily fluids were during first 4 days after birth in milliliters per kilogram per day (population mean: 127; range: 77–241).
<sup>f</sup> Data were reported only for infants in follow-up.
<sup>g</sup> Number of indomethacin doses (mean ± SD) received during prophylactic treatment: ≤3 (2.9 ± 0.4); >3 (5.7 ± 0.7) and total during hospitalization: ≤3 (2.9 ± 0.5); >3 (6.1 ± 1.0).
<sup>h</sup> Echocardiographic evidence of ductus arteriosus patency after indomethacin prophylaxis (with or without clinical symptoms).
<sup>i</sup> Data are combined echocardiographic evidence and clinical symptoms prompting indomethacin and/or ligation treatment.
<sup>j</sup> Data are neurodevelopmental impairment (cerebral palsy, MDI/FSIQ < 70, deaf, or blind) based on last assessment.
<sup>k</sup> Data are based on last assessment.
<sup>l</sup> Bayley MDI < 70 was obtained between 15 and 30 months of age.
ultrasounds (before 24 weeks’ gestation). If there were discrepancies, the early ultrasound dating was used. Necrotizing enterocolitis (NEC) was defined as Bell classification II or greater (this included NEC that was treated medically or surgically and so-called “spontaneous perforations”). Chronic lung disease (CLD) was defined as a supplemental oxygen requirement at 36 weeks’ gestational age to maintain oxygen saturation >90%. Retinopathy of prematurity (ROP) was defined as stage 2 (with plus disease) or greater than or equal to stage 3. Infants were considered to have severe ROP if they received laser treatment to at least 1 eye. Intracranial hemorrhage (ICH) was classified using the 4-level grading system. Periventricular leukomalacia was defined as echodensities that progressed to cystic degeneration. All of the infants were examined with serial bedside cranial ultrasounds initiated within the first week of life. These were repeated weekly or biweekly for the first 4 weeks. After 1 month, imaging was repeated before discharge or, more frequently, if there were any abnormal findings. A single neonatologist (Dr Clyman) prospectively evaluated and recorded all of the perinatal/neonatal risk factors and outcome measures during the hospitalization.

We also examined several long-term, neurodevelopmental outcomes. Infants surviving to nursery discharge were enrolled in a prospective nursery follow-up clinic for determination of neurodevelopmental outcome as described previously. Complete age-appropriate developmental examinations were routinely scheduled for the children at 1, 1.5, 2.5, and 4.5 years adjusted age. Age was adjusted for prematurity until 3 years. Our follow-up team traveled to several communities in Northern California to locate infants and perform assessments; infants living in these communities were examined as close to the scheduled ages as possible. Follow-up information was available for 82% of the initial population (n = 446). There were no significant differences between those who were followed and those who were lost to follow-up in gestation (followed: 25.7 ± 0.1; lost: 26.0 ± 0.1 months), male gender (followed: 55%; lost: 51%), antenatal glucocorticoid exposure (followed: 78%; lost: 76%), chorioamnionitis (followed: 22%; lost: 24%), respiratory distress (followed: 83%; lost: 89%), ICH grade 3 or 4 (followed: 4.4%; lost: 8.9%), ICH grade 3 or 4 and/or periventricular leukomalacia (followed: 7%; lost: 11%), respiratory score >3 (followed: 15%; lost: 9%), sepsis (followed: 43%; lost: 37%), NEC (followed: 12%; lost: 6%), CLD (followed: 27%; lost: 26%),ROP (followed: 15%; lost: 8%), indomethacin exposure >3 doses (followed: 33%; lost: 30%), incidence of symptomatic PDA (followed: 28%; lost: 18%), or PDA ligation (followed: 22%; lost: 17%).

A single neonatologist (Dr Piecuch), with training in developmental pediatrics, performed a neurologic examination to determine motor outcome. Final neurologic diagnosis was that obtained at the last assessment. Audiologic status was assessed by behavioral testing, and suspicious examinations were evaluated by brainstem-evoked responses or pure tone audiometry. Visual status was assessed using the near point test or Snellen eye chart. Children with questionable visual status were referred to an ophthalmologist. A single developmental psychologist (Dr Leonard) measured cognitive outcome using the Mental Developmental Index (MDI) of the Bayley Scales of Infant Development for children <36 months and the full-scale IQ (FSIQ) scores of the Wechsler Preschool and Primary Scale of Intelligence for the 4.5-year evaluation. Abnormal neurologic, neurosensory, and cognitive outcomes were defined as follows. Neurologic included moderate/severe cerebral palsy (hypotonic, spastic diplegia, hemiplegia, or quadriplegia) with functional deficits that required rehabilitative services. Neurosensory included bilateral hearing loss (requiring amplification) and blindness in either eye. Cognitive included MDI or FSIQ scores <70 (2 SDs below the mean of 100). We calculated cognitive outcome using data from the last assessment given to each child (mean age: 45 ± 23 months). We also used the Bayley MDI scores that were obtained between ages 15 and 30 months to look at the population at a narrower point in development.

**Statistical Models**

We examined the relationship among the incidence of NEC, CLD, ROP, death, and neurodevelopmental impairment with the following aspects of PDA-oriented care: (1) the presence of a symptomatic PDA, (2) the number of indomethacin doses used, (3) the ductus response to prophylactic indomethacin, and (4) the use of surgical ligation. We could not examine the effects of prophylactic indomethacin use on neonatal morbidities, because all of the infants were treated with prophylactic indomethacin. Similarly, we could not examine the effects of prolonged exposure to a symptomatic PDA on neonatal morbidities, because all of the symptomatic PDAs were closed (either pharmacologically and/or surgically) within 5 days of presentation.

We first identified the non-PDA-oriented perinatal and neonatal risk factors that were most significantly associated with the development of ROP, NEC, CLD, death, cerebral palsy, and MDI/FSIQ <70 in our population (Table 2). Estimated odds ratios (ORs) and their 95% confidence intervals were obtained for each risk factor, and a model was built for each morbidity through backward selection using P < .2 as a cutoff to keep variables in the model. We tested for interactions. These perinatal and neonatal risk factors were used in the adjusted model that evaluated the effects of PDA-oriented variables on neonatal morbidity (see model 3 below).

We next examined the effects of the different aspects
<table>
<thead>
<tr>
<th>Perinatal/Neonatal Variables</th>
<th>ROP Model</th>
<th>NEC Model</th>
<th>CLD Model</th>
<th>Death Model&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Cerebral Palsy Model</th>
<th>Cognitive Delay Model&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (95% CI)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>P</td>
<td>OR (95% CI)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>P</td>
<td>OR (95% CI)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>P</td>
</tr>
<tr>
<td>Gestation&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.58 (0.42–0.81)</td>
<td>.001</td>
<td>0.55 (0.42–0.73)</td>
<td>.001</td>
<td>0.66 (0.51–0.85)</td>
<td>.002</td>
</tr>
<tr>
<td>Year of birth 2000 or later</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
</tr>
<tr>
<td>Male gender</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
</tr>
<tr>
<td>Betamethasone exposure&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.45 (0.22–0.95)</td>
<td>.037</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
</tr>
<tr>
<td>Preeclampsia</td>
<td>3.95 (1.5–10.44)</td>
<td>.006</td>
<td>—</td>
<td>NS</td>
<td>2.05 (0.94–4.47)</td>
<td>.073</td>
</tr>
<tr>
<td>Gestational diabetes</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
</tr>
<tr>
<td>Chorioamnionitis</td>
<td>3.19 (1.48–6.86)</td>
<td>.003</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
</tr>
<tr>
<td>Respiratory distress</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
</tr>
<tr>
<td>Surfactant</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
</tr>
<tr>
<td>Respiratory score&lt;sup&gt;f&lt;/sup&gt;</td>
<td>—</td>
<td>NS</td>
<td>1.22 (1.06–1.39)</td>
<td>.005</td>
<td>1.38 (1.17–1.63)</td>
<td>.001</td>
</tr>
<tr>
<td>Intracranial hemorrhage&lt;sup&gt;g&lt;/sup&gt;</td>
<td>—</td>
<td>NS</td>
<td>2.59 (1.13–5.95)</td>
<td>.025</td>
<td>—</td>
<td>NS</td>
</tr>
<tr>
<td>Sepsis&lt;sup&gt;h&lt;/sup&gt;</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
<td>2.19 (1.24–3.84)</td>
<td>.007</td>
</tr>
<tr>
<td>Average daily fluids, mL/kg per d</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
</tr>
<tr>
<td>NEC&lt;sup&gt;i&lt;/sup&gt;</td>
<td>2.55 (1.14–5.68)</td>
<td>.023</td>
<td>NA</td>
<td>—</td>
<td>2.37 (1.13–4.97)</td>
<td>.022</td>
</tr>
<tr>
<td>Periventricular leukomalacia&lt;sup&gt;j&lt;/sup&gt;</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
<td>NA</td>
<td>—</td>
</tr>
<tr>
<td>CLD&lt;sup&lt;k&lt;/sup&gt;</td>
<td>3.97 (2.0–7.86)</td>
<td>.001</td>
<td>—</td>
<td>NS</td>
<td>NA</td>
<td>—</td>
</tr>
<tr>
<td>ROP&lt;sup&gt;h&lt;/sup&gt;</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
</tr>
<tr>
<td>Death</td>
<td>—</td>
<td>NS</td>
<td>—</td>
<td>NS</td>
<td>NA</td>
<td>—</td>
</tr>
<tr>
<td>Maternal education &lt;12 yr&lt;sup&gt;j&lt;/sup&gt;</td>
<td>NA</td>
<td>—</td>
<td>NA</td>
<td>—</td>
<td>NA</td>
<td>—</td>
</tr>
</tbody>
</table>

CI indicates confidence interval; NS, not significant; NA, not applicable; —, OR not reported because results were not significant.

<sup>a</sup>Death was during initial hospitalization.

<sup>b</sup>Cognitive delay was determined on last assessment.

<sup>c</sup>ORs are reported only for those variables that remained in the final model for the individual morbidity.

<sup>d</sup>OR is for every increase in week of gestation.

<sup>e</sup>This is defined as exposure >6 hours (combined groups 6–24 hours and >24 hours from Table 1).

<sup>f</sup>OR is for every 1 unit increase in respiratory score (mean airway pressure × fraction of inspired oxygen at 24 hours of life).

<sup>g</sup>This is defined as intraventricular hemorrhage grade 3 or higher.

<sup>h</sup>Infants with early death were not included in the calculation because outcome could not be determined.

<sup>i</sup>This is defined as average total fluids during the first 4 days of life.

<sup>j</sup>Data are reported only for infants in follow-up.
of PDA-oriented care on the development of NEC, CLD, ROP, death, and neurodevelopmental impairment. Our intent was to determine whether a particular PDA-oriented variable had an independent effect on neonatal morbidity or if its effect could be explained by immaturity, by other perinatal/neonatal risk factors, or by ≥1 of the other PDA-oriented variables. Therefore, the PDA-oriented variables were entered into 4 separate logistic regression models (model 1: each individual PDA-oriented variable by itself [the unadjusted, univariate model]; model 2: each individual PDA-oriented variable plus gestational age [a measure of immaturity]; model 3: each individual PDA-oriented variable plus the set of perinatal and neonatal risk factors identified in Table 2; and model 4: each individual PDA-oriented variable, plus gestational age, plus 1 other PDA-oriented variable).

RESULTS
Tables 3 to 5 examine the effects of the different aspects of PDA-oriented care on the development of NEC, CLD, ROP, death, and neurodevelopmental impairment. In the initial unadjusted, univariate analysis (model 1), each of the different aspects of PDA-oriented care (number of indomethacin doses used, ductus response to prophylactic indomethacin, presence of a symptomatic PDA, and use of surgical ligation) was either significantly (P < .05) or closely (P < .10) related to the development of NEC, ROP, and death (Table 3). However, once the analyses were adjusted for gestational age (model 2), there was no longer a significant association between the different aspects of PDA-oriented care and ROP, laser-treated ROP (data not shown), NEC, or death (Table 3). Therefore, the elevated risks of ROP, NEC, and death in the univariate models (model 1) could be accounted for by immature gestational age.

Similarly, in the unadjusted analysis, several aspects of PDA-oriented care were closely related to the combined long-term outcome of neurodevelopmental impairment or death (Table 4, model 1). However, once the analyses were adjusted for gestational age (model 2) or for other perinatal and neonatal risk factors (model 3), there were no longer any significant associations (Table 4). None of the other measures of long-term outcome (cerebral palsy, cognitive delay on last assessment, or cognitive delay at 15–30 months) were significantly associated with any of the PDA-oriented care variables (Table 4).

In the unadjusted model, CLD was also significantly related to the different aspects of PDA-oriented care (Table 5, model 1). However, once the analyses were adjusted for gestational age (model 2) or for other perinatal and neonatal risk factors (model 3), there were no longer any significant associations except for the use of surgical ligation (Table 5). We examined CLD as a function of the combined effects of gestational age, ligation, and each of the remaining PDA-oriented variables (see Table 5, model 4). The adverse association between each of the PDA-oriented variables and CLD completely dis-

### TABLE 3  Association Between Neonatal Morbidity During Initial Hospitalization and PDA-Related Variables (Indomethacin Exposure, Symptomatic PDA, and Ligation): Unadjusted, Adjusted for Gestation, and Adjusted for Perinatal/Neonatal Variables

<table>
<thead>
<tr>
<th>Indomethacin doses</th>
<th>PDA-Related Variables</th>
<th>ROP</th>
<th>NEC</th>
<th>Death</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (95% CI)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prophylactic doses &gt;3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted (model 1)</td>
<td>1.98 (1.08–3.62)</td>
<td>.027</td>
<td>2.08 (1.18–3.66)</td>
<td>.011</td>
</tr>
<tr>
<td>Adjusted for gestational age only (model 2)</td>
<td>1.50 (0.80–2.83)</td>
<td>.206</td>
<td>1.55 (0.80–2.80)</td>
<td>.151</td>
</tr>
<tr>
<td>Adjusted for perinatal/neonatal variables (model 3)*</td>
<td>1.38 (0.67–2.84)</td>
<td>.380</td>
<td>1.35 (0.72–2.53)</td>
<td>.351</td>
</tr>
<tr>
<td>Total doses &gt;3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted (model 1)</td>
<td>1.71 (0.95–3.06)</td>
<td>.073</td>
<td>1.79 (1.03–3.11)</td>
<td>.039</td>
</tr>
<tr>
<td>Adjusted for gestational age only (model 2)</td>
<td>1.26 (0.68–2.33)</td>
<td>.456</td>
<td>1.31 (0.73–2.34)</td>
<td>.36</td>
</tr>
<tr>
<td>Adjusted for perinatal/neonatal variables (model 3)*</td>
<td>1.35 (0.66–2.73)</td>
<td>.408</td>
<td>1.21 (0.66–2.23)</td>
<td>.539</td>
</tr>
<tr>
<td>Ductus patent after prophylactic indomethacin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted (model 1)</td>
<td>2.07 (1.00–4.27)</td>
<td>.050</td>
<td>1.99 (1.01–3.91)</td>
<td>.047</td>
</tr>
<tr>
<td>Adjusted for gestational age only (model 2)</td>
<td>1.47 (0.68–3.16)</td>
<td>.327</td>
<td>1.35 (0.66–2.77)</td>
<td>.413</td>
</tr>
<tr>
<td>Adjusted for perinatal/neonatal variables (model 3)*</td>
<td>1.33 (0.55–3.20)</td>
<td>.522</td>
<td>1.09 (0.52–2.31)</td>
<td>.816</td>
</tr>
<tr>
<td>Symptomatic PDA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted (model 1)</td>
<td>1.76 (0.93–3.32)</td>
<td>.081</td>
<td>1.78 (0.99–3.18)</td>
<td>.052</td>
</tr>
<tr>
<td>Adjusted for gestational age only (model 2)</td>
<td>1.14 (0.59–2.18)</td>
<td>.695</td>
<td>1.15 (0.62–2.15)</td>
<td>.656</td>
</tr>
<tr>
<td>Adjusted for perinatal/neonatal variables (model 3)*</td>
<td>1.24 (0.58–2.64)</td>
<td>.578</td>
<td>1.07 (0.56–2.06)</td>
<td>.830</td>
</tr>
<tr>
<td>Ligation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted (model 1)</td>
<td>1.77 (0.96–3.24)</td>
<td>.066</td>
<td>2.15 (1.19–3.91)</td>
<td>.012</td>
</tr>
<tr>
<td>Adjusted for gestational age only (model 2)</td>
<td>1.04 (0.52–2.07)</td>
<td>.918</td>
<td>1.32 (0.69–2.51)</td>
<td>.398</td>
</tr>
<tr>
<td>Adjusted for perinatal/neonatal variables (model 3)*</td>
<td>1.00 (0.45–2.21)</td>
<td>.992</td>
<td>1.19 (0.61–2.33)</td>
<td>.608</td>
</tr>
</tbody>
</table>

Cl indicates confidence interval.

* Perinatal/neonatal variables included in model 3 (see Table 2) are as follows: ROP: gestational age, chorioamnionitis, antenatal betamethasone exposure, preeclampsia, CLD, and NEC; NEC: gestational age, respiratory score, and ICH (grade 3 or greater); death (during initial hospitalization): gestational age, NEC, and ICH (grade 3 or greater).
<table>
<thead>
<tr>
<th>PDA-Related Variables</th>
<th>Neurodevelopmental Impairment or Death</th>
<th>Cerebral Palsy (Last Assessment)</th>
<th>Cognitive Delay (Last Assessment)</th>
<th>Cognitive Delay (15–30 mo)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (95% CI)</td>
<td>P</td>
<td>OR (95% CI)</td>
<td>P</td>
</tr>
<tr>
<td>Indomethacin doses</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prophylactic doses &gt;3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted (model 1)</td>
<td>2.47 (1.53–3.97)</td>
<td>.001</td>
<td>1.68 (0.54–5.21)</td>
<td>.366</td>
</tr>
<tr>
<td>Adjusted for gestational age only (model 2)</td>
<td>1.84 (1.11–3.04)</td>
<td>.018</td>
<td>1.17 (0.36–3.80)</td>
<td>.796</td>
</tr>
<tr>
<td>Adjusted for perinatal/neonatal variables (model 3)</td>
<td>1.46 (0.77–2.79)</td>
<td>.246</td>
<td>1.64 (0.47–5.74)</td>
<td>.442</td>
</tr>
<tr>
<td>Total doses &gt;3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted (model 1)</td>
<td>1.68 (0.86–3.29)</td>
<td>.129</td>
<td>1.57 (0.53–4.68)</td>
<td>.416</td>
</tr>
<tr>
<td>Adjusted for gestational age only (model 2)</td>
<td>1.48 (0.91–2.41)</td>
<td>.118</td>
<td>1.11 (0.36–3.45)</td>
<td>.859</td>
</tr>
<tr>
<td>Adjusted for perinatal/neonatal variables (model 3)</td>
<td>1.31 (0.70–2.44)</td>
<td>.401</td>
<td>1.12 (0.26–4.75)</td>
<td>.876</td>
</tr>
<tr>
<td>Ductus patent after prophylactic indomethacin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted (model 1)</td>
<td>1.98 (1.25–3.15)</td>
<td>.004</td>
<td>1.65 (0.44–6.21)</td>
<td>.458</td>
</tr>
<tr>
<td>Adjusted for gestational age only (model 2)</td>
<td>1.19 (0.59–2.42)</td>
<td>.630</td>
<td>1.12 (0.28–4.45)</td>
<td>.872</td>
</tr>
<tr>
<td>Adjusted for perinatal/neonatal variables (model 3)</td>
<td>0.98 (0.43–2.24)</td>
<td>.969</td>
<td>1.61 (0.48–5.44)</td>
<td>.445</td>
</tr>
<tr>
<td>Symptomatic PDA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted (model 1)</td>
<td>1.40 (0.79–2.46)</td>
<td>.245</td>
<td>1.40 (0.48–4.60)</td>
<td>.489</td>
</tr>
<tr>
<td>Adjusted for gestational age only (model 2)</td>
<td>0.97 (0.53–1.78)</td>
<td>.929</td>
<td>0.97 (0.30–3.18)</td>
<td>.966</td>
</tr>
<tr>
<td>Adjusted for perinatal/neonatal variables (model 3)</td>
<td>0.99 (0.49–1.99)</td>
<td>.977</td>
<td>1.06 (0.31–3.64)</td>
<td>.932</td>
</tr>
<tr>
<td>Ligation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unadjusted (model 1)</td>
<td>1.72 (0.92–3.21)</td>
<td>.087</td>
<td>1.41 (0.43–4.68)</td>
<td>.570</td>
</tr>
<tr>
<td>Adjusted for gestational age only (model 2)</td>
<td>1.14 (0.58–2.24)</td>
<td>.698</td>
<td>0.83 (0.23–2.96)</td>
<td>.772</td>
</tr>
<tr>
<td>Adjusted for perinatal/neonatal variables (model 3)</td>
<td>1.18 (0.56–2.48)</td>
<td>.671</td>
<td>0.99 (0.26–3.78)</td>
<td>.991</td>
</tr>
</tbody>
</table>

CI indicates confidence interval.

a Neurodevelopmental impairment/death indicates death (during initial hospitalization or after discharge), cerebral palsy (last assessment), MDI/FSIQ < 70 (last assessment), or bilateral hearing loss or blindness.

b Perinatal/neonatal variables included in model 3 (see Table 2) are gestational age and central nervous system injury (ICH grades 3 or 4 and/or periventricular leukomalacia).
appeared when both gestational age and surgical ligation were introduced into the models. With both gestational age and ligation in the model, the OR for the number of indomethacin doses used was only 1.02; the OR for ductus response to prophylactic indomethacin was 1.09; and the OR for the presence of a symptomatic PDA was 0.45 (Table 5, model 4). In contrast, the significant relationship between surgical ligation and CLD in the unadjusted model (model 1: OR: 2.14) was minimally affected by the introduction of either gestational age alone (model 2: OR: 1.97) or by the introduction of gestational age plus each of the other PDA-oriented variables (model 4: relationship between surgical ligation and CLD with both gestation and prophylactic indomethacin doses >3 in the model: OR: 1.79; with both gestation and total number of indomethacin doses >3: OR: 1.94; with both gestation and ductus response to prophylactic indomethacin: OR: 1.87; and with both gestation and presence of a symptomatic PDA: OR: 4.20). The independent relationship between surgical ligation and CLD was unaffected even when all of the significant perinatal and neonatal risk factors were added into the model (model 3: OR: 1.91; Table 5).

DISCUSSION

This study was designed to compare several aspects of PDA-oriented care (the presence of a symptomatic PDA, the number of indomethacin doses used, the ductus response to prophylactic indomethacin, and the use of surgical ligation) with the risk of developing NEC, CLD, ROP, death, or neurodevelopmental impairment. This study was not designed to examine the effects of prophylactic indomethacin use on neonatal morbidities, because, by design, all of the infants were treated with prophylactic indomethacin. Similarly, we could not examine the effects of prolonged exposure (>5 days) to a symptomatic PDA, because all of the symptomatic PDAs were closed, pharmacologically and/or surgically, within 5 days of presentation. Both of these variables have been examined in previous randomized, controlled trials. Indomethacin prophylaxis has been shown, both in individual studies and by meta-analysis, to have no significant impact on the incidence of NEC, CLD, ROP, death, or neurodevelopmental impairment. On the other hand, prolonged exposure to a symptomatic PDA has been shown to increase neonatal morbidity in the small number of controlled trials that have examined this issue. More recent population analyses have also found an increased incidence in neonatal morbidity after prolonged exposure to a symptomatic PDA.

Because infants were not allowed to have a symptomatic PDA for >5 days, we were able to examine the effects of other PDA-oriented variables without the confounding effects of prolonged exposure to a symptomatic PDA in the analysis. We found that neither the ductus response to prophylactic indomethacin nor a brief (<5-day) exposure to a symptomatic PDA led to the subsequent development of ROP, NEC, CLD, or death. The apparent morbid effects of these 2 variables, observed in the unadjusted models (model 1), became insignificant when the models were adjusted for gestational age (Tables 3 and 5, model 2). Developmental processes determine the incidence of PDA, its response to indomethacin, and the incidence of ROP, NEC, and CLD. It is not surprising that morbidities that are similarly caused by interrupted development seem to be associated with each other in univariate analyses.

Similarly, we found that, when gestational age was included in the analysis, the number of indomethacin doses that an infant received (either during the prophylactic treatment course or during the entire hospitalization) was not significantly associated with the incidence of ROP, NEC, or death (Table 3, model 2). The apparent adverse effect of the number of indomethacin doses on the incidence of CLD could be completely explained by the presence of both gestational age and surgical ligation (Table 5, model 4). This is consistent with the results of controlled trials of short versus long courses of indomethacin.

---

**TABLE 5**  
Association Between CLD and PDA-Related Variables: Unadjusted, Adjusted for Gestation, Adjusted for Perinatal/Neonatal Variables, and Adjusted for Gestation and Ligation

<table>
<thead>
<tr>
<th>PDA-Related Variables</th>
<th>Model 1: Unadjusted OR (95% CI)</th>
<th>Model 2: Adjusted for Gestational Age, OR (95% CI)</th>
<th>Model 3: Adjusted for Perinatal and Neonatal Factors, OR (95% CI)</th>
<th>Model 4: Adjusted for Gestational Age and Ligation, OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indomethacin doses</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prophylactic doses &gt;3</td>
<td>2.09 (1.26–3.47)b</td>
<td>1.69 (1.00–2.86)</td>
<td>1.35 (0.75–2.44)</td>
<td>1.32 (0.71–2.45)</td>
</tr>
<tr>
<td>Total doses &gt;3</td>
<td>1.83 (1.13–2.95)b</td>
<td>1.44 (0.87–2.38)</td>
<td>1.23 (0.70–2.16)</td>
<td>1.02 (0.54–1.94)</td>
</tr>
<tr>
<td>Ductus patent after prophylactic indomethacin</td>
<td>2.33 (1.25–4.36)b</td>
<td>1.79 (0.93–3.45)</td>
<td>1.54 (0.75–3.18)</td>
<td>1.09 (0.44–2.70)</td>
</tr>
<tr>
<td>Symptomatic PDA</td>
<td>2.81 (1.65–4.78)b</td>
<td>1.54 (0.90–2.64)</td>
<td>1.55 (0.85–2.81)</td>
<td>0.45 (0.10–2.06)</td>
</tr>
<tr>
<td>Ligation</td>
<td>2.14 (1.29–3.55)b</td>
<td>1.97 (1.11–3.47)b</td>
<td>1.91 (1.02–3.57)</td>
<td>—</td>
</tr>
</tbody>
</table>

CI indicates confidence interval.

a Perinatal/neonatal variables included in model 3 of CLD (see Table 2) included the following: gestational age, male gender, preeclampsia, respiratory score, NEC, and sepsis.

b P < .05.
methacin treatment, where no relationship was found between the duration of indomethacin exposure and the incidence of ROP, NEC, CLD, or death.\textsuperscript{15}

Surgery, during the neonatal period, has been associated with neurodevelopmental problems after preterm birth.\textsuperscript{18–40} Our findings are somewhat reassuring, because we found no relationship between ductus ligation (or any of the other ductus-oriented factors) and subsequent neurodevelopmental problems (Table 4). On the other hand, surgical ligation was found to be significantly and independently associated with an increased incidence of CLD. The relationship between ligation and CLD was independent of an infant’s gestational age (Table 5, model 2) or of other PDA-related variables (model 4). Surgical ligation was still a significant, independent risk factor for CLD, even when other perinatal and neonatal risk factors for CLD were incorporated into the model (Table 5, model 3).

Our study cannot determine whether surgical ligation plays a causative role in the development of CLD or is simply a surrogate marker for infants with a unique developmental profile leading to CLD. Recent findings in premature baboons support the concept that surgical ligation may have a detrimental effect on lung function and growth. Premature newborn baboons, exposed to a moderate-size PDA shunt for 2 weeks, have decreased pulmonary function and arrested alveolar development.\textsuperscript{41} Pharmacologic closure of the PDA prevented the deterioration in both pulmonary function and alveolar development.\textsuperscript{41} In contrast, surgical closure of the PDA offered no benefit for either pulmonary function or for alveolar growth or development.\textsuperscript{42} Similarly, in the only published controlled trial to compare pharmacologic closure of the PDA with surgical ligation,\textsuperscript{19} infants who were surgically ligated tended to need longer durations of continuous positive airway pressure than those treated with indomethacin ($P = .06$). In sum, these studies suggest that ductus ligation, while eliminating 1 potential cause for neonatal morbidity, may introduce its own set of problems.

Our findings add to the growing uncertainty about the benefits and risks of surgical ligation\textsuperscript{1} of or surgery in general during the neonatal period.\textsuperscript{18–40} Additional investigations will be needed to determine which infants are most likely to benefit from surgical ligation of their PDA and which infants might best be left untreated.

ACKNOWLEDGMENTS

This research was supported by National Institutes of Health grants HL466911 and HL56061 and by a gift from the J. and B. Gates Foundation. This study was carried out in part in the Pediatric Clinical Research Center (Moffitt Hospital, University of California, San Francisco) with funds provided by National Center for Research Resources grant 5 M01 RR-01271, US Public Health Service.

We thank the fellows and attending staff of the Division of Pediatric Cardiology, who have been so helpful in performing and interpreting the echocardiographic studies, and the Intensive Care Nursery and Pediatric Clinical Research Center nurses, without whom this study would not have been possible. Dr Chuck McCulloch provided invaluable statistical advice and support.

REFERENCES


Continuity of Health Insurance Coverage Among Young Adults With Disabilities

S. Todd Callahan, MD, MPH, William O. Cooper, MD, MPH

Divisions of Adolescent Medicine and General Pediatrics, Child and Adolescent Health Research Unit, Department of Pediatrics, Vanderbilt University Medical Center, Nashville, Tennessee

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

ABSTRACT

OBJECTIVES. Although considered critical to facilitate the successful transition from pediatric to adult health care, the continuity of insurance coverage for young adults with disabilities as they make the transition to adulthood has not been well characterized. The purpose of this work was to compare the continuity of insurance coverage reported by a nationally representative sample of young adults 16 to 25 years old with and without disabilities during a consecutive 36-month period.

METHODS. We performed secondary analysis of data from the 2001 Survey of Income Program and Participation. Data for the survey were collected at 4-month intervals from February 2001 through January 2004 for 5170 young adults. Subjects with disabilities were those who reported limitations in activities of daily living or work, used assistive devices, and/or had learning disabilities, mental retardation, or other mental disorders. The primary outcome was uninsurance defined for each study month in which there was no coverage by private, public, or military programs. We present bivariate analyses of the months of uninsurance according to disability status using longitudinal weights and design-effect adjustments to account for the complex sample design.

RESULTS. The sample consisted of 599 subjects with and 4571 without reported disabilities, representing 3,970,000 and 30,800,000 young adults in the United States, respectively. At study entry, 22% of the young adults with disabilities were uninsured. During the 36-month follow-up period, 56% of the young adults with disability reported gaps in insurance coverage with a mean of 15 months of uninsurance. The proportion of uninsured subjects did not significantly differ by disability status.

CONCLUSIONS. The majority of young adults with disabilities reported gaps in insurance coverage, and many were uninsured for a substantial portion of the study period. As an increasing number of children with special health care needs make the transition to adulthood, improving the continuity of health insurance coverage for this population warrants specific attention.
Because an estimated 500 000 children with disabilities and other special health care needs reach adulthood each year, the health care transitions of these youth have garnered increasing attention. Among the many life and health care transitions that characterize this period, young adults must also negotiate changes in health insurance coverage. Many of these youth have ongoing health issues; therefore, continuous health insurance coverage is desirable to allow timely access to health care during young adulthood and to facilitate the transition to adult health care.

The continuity of health insurance coverage for young adults with disabilities has not been well characterized to date. Many become ineligible for the public coverage or their parent’s private health insurance at age 19 years. Full-time students may be eligible for continued coverage on their parents’ policy until they complete college. However, as they age out of the coverage that they held as children, young adults are twice as likely to be uninsured as children or older adults, and studies suggest that young adults with disabilities have rates of uninsurance that are similarly high. Most current studies of health insurance coverage for young adults with disabilities are limited to cross-sectional data that fail to capture discontinuities in health insurance coverage over time and may underestimate the numbers who are uninsured during young adulthood.

The 3 specific aims of this study were to use longitudinal survey data to compare the continuity of health insurance coverage and the number of months of uninsurance reported by adolescents and young adults with and without disabilities during a 36-month period and to assess the continuity of health insurance for different age groups of young adults with disabilities.

METHODS

Data Source

We performed secondary analyses of data from the 2001 Survey of Income Program and Participation (SIPP). The SIPP is a federally funded longitudinal multiyear panel survey of the civilian noninstitutionalized population administered by the US Census Bureau. The purpose of the SIPP is to collect information on income, labor force, and program participation of individuals and households in the United States. The 2001 SIPP included in-person and telephone interviews conducted at ~4-month intervals from February 2001 through January 2004 for a total of 9 interviews over 36 months. Because of budget constraint, the US Census Bureau cut the size of the survey sample by ~25% between waves 1 and 2 (from 40 500 eligible households to 30 500 eligible households). The sample was not cut for the remaining waves (waves 3–9). Original sample members were followed regardless of subsequent changes in residence. The US Census Bureau used imputation procedures to handle missing data. Longitudinal weights were provided by the US Census Bureau to account for panel attrition and poststratification adjustments to make the weighted sample totals conform to population totals for key variables. The survey weights permit estimations that project to the noninstitutionalized population in the United States. SIPP data are publicly available as microdata files and may be downloaded from www.sipp.census.gov.

Study Population

We included all of the respondents who were 16 to 25 years old at the time of the first interview and for whom information was available for the entire 36 months of the survey. We chose to include respondents with the entire 36 months of data to allow for a longitudinal assessment of insurance status for longer periods of time. Of the 9234 young adults who completed wave 2 of the 2001 SIPP, 5170 subjects (56%) had complete data for the entire 36 months.

Definition of Disability

Respondents were asked about physical and mental limitations during interviews conducted at months 8, 20, and 32. Using survey data, we defined disability as a reported learning or mental impairment (including learning or developmental disability and mental retardation) or limitations in activities of daily living (ADL), instrumental activities of daily living (IADL), or work because of a physical or mental condition and/or the use of assistive devices, such as hearing aids or canes for ≥6 months. ADL included getting around inside the home, getting in or out of a bed or chair, bathing, dressing, and toileting. IADL included getting around outside the home, taking care of money and bills, preparing meals, doing light housework, or using the telephone. This definition of disability has been used previously to estimate rates of disability in the United States.

The comparison group included 16- to 25-year-olds without reported limitation. Based on cell size requirements published by the US Census Bureau, the sample was sufficient to provide reliable national estimates of the entire US population of young adults with and without disabilities.

Health Insurance Coverage and Sociodemographic Variables

The primary outcome was health insurance coverage, which was determined for each study month and expressed as a dichotomous variable. Respondents were considered insured for every month that they reported coverage by private, public, or military programs. Continuously insured young adults reported insurance coverage for each month of the 36-month period, whereas those with gaps reported ≥1 month of noncoverage.

We used survey data to identify the respondent’s age at study entry, gender, race, ethnicity, employment and
student status, and household income. Young adults often age out of public and private coverage held as children at age 19 or 22 years; therefore, we stratified young adults into 3 age categories based on their age at the initial interview: 16 to 18 years, 19 to 21 years, and 22 to 25 years. Household income was expressed as a percentage of the US Census Bureau’s poverty threshold for each survey year after considering overall family size and number of children in the household. For reference, the weighted average poverty threshold for a family of 4 was $18,104 in 2001 and $19,307 in 2004.

Data Analysis
We used design effect adjustments provided by the US Census Bureau to account for the complex sample design and longitudinal weights to project the sample to the US noninstitutionalized population. We calculated SEs of the estimates of sociodemographic and insurance characteristics for young adults with and without disabilities using generalized variance parameters provided by the US Census Bureau. To test for differences in the characteristics of young adults with and without disabilities, we calculated SEs of the difference of sample estimates using formulae provided by the US Census Bureau. We report significance at the 5% level when the estimates differed by \(1.96\) times the calculated SE of the difference and at the 1% level when estimates differed by \(2.57\) times the calculated SE of the difference. The study was approved by the institutional review board of Vanderbilt University Medical Center.

RESULTS
In the study population, 11.4% of 16- to 25-year-olds reported having a disability, representing ~4 million young adults with disabilities in the United States. Among those with disabilities, 48% had learning or mental impairments, 45% were limited in functional activities, 37% reported limitations in the ability to work, 16% had limitations in ADL or IADL, and 7% used assistive devices. Overall, 58% of young adults with disabilities reported being limited in >1 of these disability domains. Young adults with disabilities did not differ from peers without disabilities in terms of age, Hispanic ethnicity, or marital status at the time of the initial interview (Table 1). Young adults with disabilities were significantly more likely than peers without disabilities to be black and to live in a household with an income <200% of the poverty threshold and were less likely to be attending school or to be employed at the time of the initial interview.

Insurance Coverage According to Disability Status
Table 2 displays the health insurance characteristics of young adults with and without disabilities. Young adults in both groups were equally likely to be insured at the start of the study. Compared with peers without disability, young adults with disabilities were significantly less likely to have private health insurance \((P < .01)\) and significantly more likely to have public coverage at the start of the study \((P < .01)\).

Continuity of Insurance Coverage According to Disability Status
The majority of young adults with disabilities reported gaps in health insurance coverage, and young adults with disabilities were as likely as those without disabilities to report having a gap in health insurance coverage (56% vs 54%). Twenty-eight percent of young adults

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Reported Disability</th>
<th>No Disability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male, % (SE)</td>
<td>48.9 (2.7)</td>
<td>48.2 (0.9)</td>
</tr>
<tr>
<td>Age, % (SE)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16–18 y</td>
<td>31.6 (2.5)</td>
<td>32.2 (0.9)</td>
</tr>
<tr>
<td>19–21 y</td>
<td>26.9 (2.4)</td>
<td>30.6 (0.9)</td>
</tr>
<tr>
<td>22–25 y</td>
<td>41.5 (2.6)</td>
<td>37.2 (0.9)</td>
</tr>
<tr>
<td>Race, % (SE)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic white</td>
<td>66.8 (2.5)</td>
<td>66.3 (0.9)</td>
</tr>
<tr>
<td>Non-Hispanic black</td>
<td>16.9 (2.0)</td>
<td>12.8 (0.6)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>12.3 (1.7)</td>
<td>15.1 (0.7)</td>
</tr>
<tr>
<td>Other</td>
<td>4.0 (1.0)</td>
<td>5.8 (0.5)</td>
</tr>
<tr>
<td>Married, % (SE)</td>
<td>16.2 (2.0)</td>
<td>16.8 (0.7)</td>
</tr>
<tr>
<td>Household income (percentage of poverty threshold), % (SE)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0%–100%</td>
<td>14.9 (1.6)</td>
<td>8.8 (0.5)</td>
</tr>
<tr>
<td>101%–199%</td>
<td>24.6 (2.0)</td>
<td>18.8 (0.6)</td>
</tr>
<tr>
<td>≥200%</td>
<td>60.6 (2.2)</td>
<td>72.4 (0.7)</td>
</tr>
<tr>
<td>Enrolled in school (full- or part-time), % (SE)</td>
<td>44.2 (2.6)</td>
<td>54.5 (1.0)</td>
</tr>
<tr>
<td>Holds a paid job, % (SE)</td>
<td>58.5 (2.6)</td>
<td>69.3 (0.9)</td>
</tr>
</tbody>
</table>

Data source was our analysis of the 2001 SIPP. Data are the difference between groups significant at \(P < .05\).
with disabilities were uninsured between 1 and 12 months, and another 28% were uninsured for ≥13 of the 36 months of the study period. Youth with disabilities who experienced gaps in coverage reported being uninsured for a mean of 15 months out of the 36-month period. The mean number of months without health insurance did not significantly differ for young adults with or without disabilities.

### Insurance Coverage of Young Adults With Disabilities According to Age

When stratified by age at the start of the study, youth who were 16 to 18 years old were significantly less likely than older peers to be uninsured at the start of the study. They were also less likely to report a gap in coverage during the 36-month period \( P < .05 \). As shown in the Fig 1, 5% of 16- to 18-year-olds were uninsured at the start of the study, and 46% reported a gap in coverage during the next 36 months. Among 19- to 21-year-olds, 30% were uninsured at the start of the study, and 62% reported a gap in coverage during the study period. Similar rates were seen for 22- to 25-year-olds, with 30% reporting uninsurance at the start of the study and 60% reporting a gap in coverage during the study period. Although 18% of 16- to 18-year-olds were uninsured for >12 months, one third of 19- to 21- and 22- to 25-year-olds were uninsured for >12 months.

### DISCUSSION

In this study, we used longitudinal data to assess characteristics of insurance coverage during a 36-month period for a nationally representative sample of young adults with and without disabilities. We found that the majority of young adults in our study experienced a gap in health insurance coverage during the study, and many were uninsured for a substantial period of time. The proportion of young adults with disabilities who experienced gaps in coverage and the mean number of months of uninsurance did not differ from the group of youth without disabilities.

Previous studies using cross-sectional data provide estimates that 20% to 26% of young adults with disabilities are uninsured at a point in time.\(^8^{12}^{13}\) The findings of our study add to the literature by using longitudinal data.
to assess insurance status over time. Although the proportion of young adults who were uninsured at the start of the study is consistent with previous point-in-time estimates, the longitudinal data show that 56% of young adults with disabilities were uninsured during the 3-year study. Because they are more likely to have ongoing health issues, we hypothesized that young adults with disabilities would have greater incentive to maintain continuous health insurance coverage than their peers without disability. Our study findings did not support this hypothesis but rather suggest that the chasm between health insurance coverage options for children with disabilities and those for adults creates a period of particular vulnerability for young adults with disability.

In our study, young adults over age 18 years at the start of the study were more likely than those less than age 18 years to experience gaps in coverage. Most children with disabilities are insured through public coverage or as a dependent on their parent’s private health insurance, and the majority will lose eligibility for both of these types of coverage during young adulthood. Medicaid coverage is not available to most young adults unless they are parenting. Supplemental Security Income provides coverage for only one fourth of children with disabilities, and requirements become more stringent at age 19 years. Consequently, approximately one third of adolescents who receive Supplemental Security Income do not qualify for similar coverage in adulthood. Young adults with disabilities may be less likely to have access to private health insurance than peers without disability as demonstrated by the differences in insurance coverage at baseline in our study. The majority of adults obtain private health insurance through employment-based coverage. Youth with disabilities are less likely to be employed or to work full-time than peers without disability; thus, it is likely that they would have greater difficulty procuring employment-based health insurance coverage.

Our study also demonstrates that youth with discontinuous health insurance coverage are uninsured for extended periods of time. We found that young adults with disabilities who experienced gaps in coverage were uninsured for a mean of 15 of the 36 months. Many youth with disabilities require ongoing health care to reduce morbidity and improve functioning. As young adults with disability transfer from pediatric to adult health care providers, lack of health insurance may limit their options for adult health care. In addition to its effects on access to health care for disability management, lack of health insurance is likely to reduce access to preventive health care.

We anticipated that a greater need and awareness for health care might positively affect insurance rates for young adults with disabilities relative to those without disabilities. However, it is also possible that persons with disabilities may have less opportunity to gain health insurance through employment or full-time student status, the means by which most young adults are covered. Thus, it is also plausible that disability status could result in lower rates of insurance and higher rates of discontinuity. These social determinants of health and health care access are important issues, and additional study to assess the potential health impact of discontinuities in health insurance in this population will be important.

Some states have proposed or enacted policies that may reduce gaps in health insurance coverage during young adulthood. Several states have enacted or are considering legislation to raise the age limit for dependent health insurance coverage, including New Jersey, which passed legislation in 2006 that allows young adults to qualify for dependent health insurance coverage up to age 30 years. Increasing the age of eligibility for dependent coverage may benefit young adults with disabilities, because many of these youth will not qualify for public coverage based on their disability and/or income. In recent years, New Jersey, Massachusetts, and other states extended eligibility for public health insurance programs like Medicaid and State Children’s Health Insurance Program beyond age 18 years to childless young adults who meet income guidelines. Unfortunately, provisions in these states were subsequently scaled back because of budgetary issues. Pediatric and young-adult providers should keep abreast of initiatives in their state and consider contributing their expertise to local and national policy initiatives.

This study has several potential limitations. The SIPP is based on self-report and may be subject to recall and nonresponse bias. However, imputation procedures and the brief time between interviews are designed to minimize these sources of bias. In addition, nonresponse and loss to follow-up may make the study sample differ from the underlying population. Although weighting procedures and sampling techniques used by the US Census Bureau account for declining participation, this potential limitation is important to take into consideration in interpreting the study results. Assessment for disability occurred at differing points during study follow-up; therefore, we do not know how many of the young adults developed limitations after the onset of the study or if those limitations resolved within the study period.

Despite increasing focus on issues of health care transition for young adults with disabilities, this study suggests that the majority of youth with disabilities will be uninsured during young adulthood, many for a substantial period of time. Uninsurance rates for young adults with disabilities were comparable to young adults without disabilities. Expanding the availability of private and public coverage and developing other strategies to reduce gaps in coverage are potentially important targets to improve health care for young adults with disabilities. Policy changes suggested by others have included extending the age for dependant coverage through private
insurance through the age of 23 years, extending eligibility for Medicaid and other government programs through the age of 23 years, and ensuring that colleges and universities require all students to have some form of insurance coverage. Additional research should examine the implications of uninsurance on the health and health care access of young adults with disabilities.

ACKNOWLEDGMENT

This study was supported by a grant from the Vanderbilt Physician Scientist Development program (to Dr Callahan).

REFERENCES

Neonatal Intensive Care Unit Oxygen Management: A Team Effort

Jay P. Goldsmith, MD\textsuperscript{a}, Jay S. Greenspan, MD, MBA\textsuperscript{b}

\textsuperscript{a}Department of Pediatrics, Tulane University, New Orleans, Louisiana; \textsuperscript{b}Departments of Pediatrics and Neonatology, Thomas Jefferson University and Alfred I. duPont Hospital for Children, Philadelphia, Pennsylvania

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

"Because I said so!" We have all been guilty of using this abrupt paternalistic response with our children when they question a direction. Maybe at the time we were too busy to give a full explanation, but the result may have been that the compliance was as brief as the response. Perhaps a more reasoned explanation would bring a better and more lasting result.

It is no different when we as physicians give orders to our health care team. Orders without education or explanation, especially when they differ from previously accepted norms, will result in less compliance than when the health care team is fully aware of the reasons for the changes and “buys in” to the new approach.

This principle of quality improvement is well illustrated by the use of oxygen in our neonatal intensive care units (NICUs). Bedside neonatal caregivers see oxygen as life-giving, not poisonous. Establishing new oxygen targeting for preterm infants has become a hot issue in neonatology. It is well recognized that oxygen toxicity leads to a multitude of complications, and reducing the levels and time of oxygen exposure to sick infants from the delivery room through their early intensive care course will likely reduce morbidities. However, allowing very low birth weight infants to live in previously undesirable oxygen saturation ranges may be a difficult culture to change.

The challenges of translating the concept of lower oxygen targets to practice have many obstacles. In this issue of \textit{Pediatrics}, Clucas et al\textsuperscript{1} document that only 22\% of their population of preterm infants had pulse oximetry alarms set in accordance with the established protocol. Even when the saturation limits are properly set, Hagadorn et al\textsuperscript{2,3} have shown previously that without a special compliance program, they were able to keep sick oxygen-dependent infants in their 14 esteemed NICUs within the targeted range only about half of the time.

Clucas et al reported that a memo sent to all staff citing the new standard for oximetry limits in the neonatal unit of the Royal Woman’s Hospital resulted in a low compliance even when the nursery team knew that they were being monitored. It is unclear why the orders were not followed very often. Were the guidelines clear, well dispersed, and reinforced? Did the staff know the targets but choose to ignore them? What educational process was given to the caretakers? We know that an intensive education program for hand-washing improves compliance, but still only about half of the health care team performs this task correctly, and the education must be repeated frequently to maintain even that much compliance.\textsuperscript{4}

Certainly clinical leadership should view this study with great concern. When a change in clinical culture is necessary, the process requires careful planning and execution. In the case of saturation limits, it is likely that frequent alarms may prompt a caretaker to reset limits, especially if the ordered limits target too narrow a range. It is possible that the care team ignored the guideline or that there was difficulty with the equipment. Unfortunately, the most critically ill infants, and the infants who

Abbreviation: ROP, retinopathy of prematurity

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.2007-0462
doi:10.1542/peds.2007-0462

Accepted for publication Feb 16, 2007
Address correspondence to Jay P. Goldsmith, MD, 1625 Joseph St, New Orleans, LA 70115. E-mail: goldsmith.jay@gmail.com

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
would likely benefit most from the lower targeted saturations that Clucas et al described, are the most labile and, therefore, the most difficult to keep within range.1

In the complex NICU environment, creating a major change, such as resetting the acceptable oxygen saturation target ranges for sick neonates, involves many obstacles. Such change requires a multidisciplinary team to develop the vision and strategy, assess and eliminate roadblocks, establish a sense of urgency, and communicate the new guidelines to all parties involved. There should be short-term wins, probably with rewards, a process of assessment and quality improvement, and reinforcement until the change becomes anchored as part of the culture. The process may uncover the need for new equipment (ie, different monitors), alterations in staffing models (ie, 1:1 nursing for the smallest and more labile infants), evidence-based education for staff and/or families, etc.

There are several published demonstrations that lower oxygen targets can be achieved and resultant morbidities such as retinopathy of prematurity (ROP) can be reduced.5–7 Chow et al5 used contracts with the nursing and respiratory staffs and the establishment of the Caring Responsible Approach to Development in the Lives of Extremely Low Birth Weight Infants (CRADLE) Club to promote compliance with the new targeted ranges and achieved virtual elimination of surgical intervention for ROP in their unit over 3 years. However, no measures were made of how much time the infants spent in the targeted ranges. The Oxygen With Love (OWL) program uses an icon on each bed (an owl, recognized for wisdom and good eyesight) to remind staff of the targeted ranges to supplement unit education, which includes the parents. Monitoring targeted ranges in this unit resulted in 80% compliance with a similar reduction in ROP for inborn infants (J.P.G., unpublished data).

There is general agreement that oxygen targets for critically ill preterm infants should be lower than historic norms. Clucas and colleagues have demonstrated that the first step in achieving this change, altering the alarm limits, is difficult to achieve. Education, involvement of the entire health care team and the parents, frequent monitoring and assessment with broad-based communication, empowerment, and perseverance will be needed to reach the goal of optimal oxygen delivery for sick neonates.

REFERENCES
Drug Pricing in Pediatrics: The Egregious Example of Indomethacin

Alan H. Jobe, MD, PhD

Division of Pulmonary Biology, Cincinnati Children’s Hospital, Cincinnati, Ohio

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

Indomethacin for closure of a patent ductus arteriosus (PDA) in preterm infants has been the standard of care since the 1970s. Indomethacin was the first drug specifically approved for use in infants by the US Food and Drug Administration as the result of collaborations between academic pediatricians and the pharmaceutical company Merck. Indomethacin was not a new chemical entity when evaluated for PDA, and it remains today one of the most frequently used over-the-counter nonsteroidal antiinflammatory drugs. The unique aspect of indomethacin for PDA is that the drug is formulated for intravenous use. There are no other approved uses for intravenous indomethacin in the United States.

Ovation Pharmaceuticals acquired the distribution rights for indomethacin for PDA from Merck in 2006. The list price for indomethacin then increased from approximately $100 to $1875 for three 1-mg vials. This is a rather astounding increase in price for a drug that has a stable niche market and requires no advertising, no educational expenses (all neonatologists know how to use indomethacin), and no further drug development. It is quite hard to imagine how such an increase in price could be justified. This concern, together with the public discussions about Medicare negotiations for drug prices, stimulated me to ask colleagues about indomethacin pricing in other countries. Hospital costs of indomethacin in available packaging and the cost per milligram of drug in US dollars are shown in Table 1 for comparative purposes. My unease about the pricing in the US was further strengthened. The price per milligram is ∼30 to 60 times higher in the United States than in other countries with health care systems of similar quality. I suspect that manufacturers and distributors in these other countries are not losing money.

I also have another concern. There are a number of recent reports that indicate that ibuprofen is as effective as indomethacin for the closure of PDA in preterm infants and it may have a better safety profile.1,2 The initial dose of ibuprofen is 10 mg/kg or 50 times the dose of 0.2 mg/kg for indomethacin. Ibuprofen is also commonly used as an over-the-counter nonsteroidal antiinflammatory drug. An intravenous formulation of ibuprofen was recently approved by the US Food and Drug Administration only for use for closure of PDA in preterm infants. Ibuprofen is less widely available for this indication worldwide, but it is ∼10 times more expensive in the US than in the United Kingdom or Germany (Table 1). Ovation Pharmaceuticals is also the only source of ibuprofen for closure of PDA in the United States, and

### Table 1: Comparative Pricing

<table>
<thead>
<tr>
<th></th>
<th>Indomethacin</th>
<th>Ibuprofen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit Size, mg</td>
<td>$/mg (Total Cost, $)</td>
<td>Unit Size, mg $/mg (Total Cost, $)</td>
</tr>
<tr>
<td>United States</td>
<td>3 × 1 (1875) 613</td>
<td>3 × 20 (1812) 30</td>
</tr>
<tr>
<td>Canada</td>
<td>1 (14) 14</td>
<td>— —</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>3 × 1 (47) 16</td>
<td>4 × 10 (522) 3</td>
</tr>
<tr>
<td>Germany</td>
<td>12 × 1 (264) 22</td>
<td>4 × 10 (522) 3</td>
</tr>
<tr>
<td>Holland</td>
<td>1 (22) 22</td>
<td>— —</td>
</tr>
<tr>
<td>Australia</td>
<td>3 × 1 (33) 11</td>
<td>5 (130) 26</td>
</tr>
</tbody>
</table>

— indicates not available in country.

Abbreviation: PDA, patent ductus arteriosus

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.2007-0184
doi:10.1542/peds.2007-0184

Accepted for publication Jan 24, 2007
Address correspondence to Alan H. Jobe, MD, PhD, Cincinnati Children’s Hospital, Division of Pulmonary Biology, 3333 Burnet Ave, Cincinnati, OH 45229-3039. E-mail: alan.jobe@cchmc.org

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics

PEDIATRICS Volume 119, Number 6, June 2007 1197
the retail price for a treatment course of three 20-mg vials is $1812. This drug does need further evaluation, but its costs are just $53 less than indomethacin. To the clinical neonatologist or pharmacist, the similar pricing of ibuprofen and indomethacin seems to be more than a coincidence. We all know that drugs are expensive in the United States and that we subsidize drug development for the rest of the world. However, the pricing of these 2 useful agents that are standard of care is quite extraordinary. Words such as “unconscionable,” “unethical,” and “socially irresponsible” come to mind.

REFERENCES

GENETIC TESTS OFFER PROMISE, BUT RAISE QUESTIONS, TOO

“A growing industry is hoping to spin gold from DNA’s double helixes by using ultra-sensitive genetic tests to personalize medical treatment for cancer, lupus and other diseases. . . . More than 1000 genetic tests are available to researchers. Despite the tremendous promise of these tests, there is growing concern among researchers and patient advocates about how consistently their claims match reality. How accurate are they at finding potential genetic problems? Are different tests for different conditions equally reliable? And how tight is the connection between a genetic trait and a specific illness? Some researchers say they believe that the practical relevance of many tests has been oversold. Over the last two decades, for example, there has been a steady stream of news about researchers discovering ‘the gene’ that links people to diabetes, Alzheimer’s, obesity, schizophrenia, depression and many other afflictions. Yet most of those hard-wired gene-disease links—as many as 95 percent of them, according to one British study published in 2003—don’t hold up to closer scrutiny. Instead, follow-up studies find that if there is any measurable genetic link to these common diseases, it results from the more complex interactions of many genes with one another, as well as with the environment. According to the Human Genome Project, this state of affairs is particularly troubling, considering that a few companies have started marketing genetic tests directly to the public—sometimes claiming their kits not only test for disease, but can also customize medicine, vitamins and diet to an individual’s genetic makeup. There is no independent review or government oversight of the validity of these tests, particularly those available to consumers through their doctors. No agency yet has the formal responsibility to make sure that genetic tests can produce correct answers reliably over time, or, more important, that there is even a relationship between a particular genetic variation and a person’s health.”


Noted by JFL, MD
COMMENTARY

Counseling Youth About Military Service Options and Selective Service Registration: An Integral Part of Anticipatory Guidance of Adolescents

Carl E. Stafstrom, MD, PhD

Section of Pediatric Neurology, Departments of Neurology and Pediatrics, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin

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

Pediatricians play a crucial role in counseling adolescents about numerous aspects of health and well-being. Anticipatory guidance for this age group includes discussion of risk-taking behaviors; drug, tobacco, and alcohol use; depression and suicide; sexuality and safe sex practices; family planning and dynamics; eating disorders; and a wide variety of other health and developmentally related topics.1-4 Counseling this age group may also include issues related to planning for higher education and future vocation. I propose that anticipatory guidance of adolescents should also include discussion about military service and alternatives, the laws regarding registration (and consequences for failure to register) with the Selective Service System (SSS), and the medical and psychological effects of war. Specific developmental challenges face adolescents who opt for (United States) or are required to participate in (many other countries) military service. By late adolescence, individuals have usually developed the ability to think in abstract terms, plan for their future, and engage in independent life choices,5 yet this age group still lacks full psychosocial maturation. Exposure to the demands of military service (absolute conformity and obedience, lack of independence, separation from family, real risk of physical harm) places unique stresses on adolescents in peacetime and even more so during times of war.3

The purpose of this commentary is to raise awareness among pediatricians and other providers of adolescent health care about current SSS laws, possible changes to such laws that could occur at any time, and options available to adolescents who, by reason of conscience, choose not to participate in military service or perhaps even register with the SSS. To my knowledge, this topic is rarely if ever discussed in the context of anticipatory guidance of adolescents facing educational and career choices, and extensive search of the literature failed to reveal any references on this topic. Providing adolescents with up-to-date information about SSS requirements will allow them to make an informed decision about registration according to their personal beliefs. Although my suggestion, to add discussion of military service issues to anticipatory guidance, requires additional time and effort on the part of health care providers, the timeliness and importance of this information cannot be understated.

WHAT IS THE SSS?
The SSS is a US government agency charged with organizing a draft should one be mandated by the President and authorized by Congress. The current task of the SSS is to maintain an updated registry of men aged 18 to 25 years, from which men can be drafted to supplement the current all-volunteer armed forces.6 In the most recent draft, during the Vietnam War, young men were conscripted according to a lottery on the basis of their birth date; after medical screening, each man reported to his local Selective Service board (“draft board”). Each municipality maintains a draft board. In the case of an active draft, the local draft board is responsible for classifying SSS registrants as to their eligibility and suitability for...
military service and deciding whether the individual qualifies for exemption, postponement, or deferment of military service (see www.sss.gov). Today, draft boards still exist but operate on standby status only, because there is no active draft. However, draft board members still prepare for the possibility of a new draft with annual training sessions.7

WHAT IS THE CURRENT SELECTIVE SERVICE LAW?
There has not been a draft since 1973, but the law still requires every man living in the United States to register with the SSS within 30 days of his 18th birthday. Ironically, even illegal immigrants (other than those admitted on immigrant visas) are required by law to register with the SSS. Registration is accomplished by filling out a card that will arrive in the mail to some eligible men; the form is also available at any post office or online (at www.sss.gov). The SSS keeps registrant information in a database so that, should a draft be called, there is a pool of individuals readily available (currently estimated at 16 million men). Repercussions for failing to register with the SSS could include a $250,000 fine and/or a jail term of up to 5 years. It is important to note that there have been no prosecutions for failure to register with the SSS since 1985. However, 41 states and Washington, DC have laws that deny obtaining (or renewing) a driver’s license and state loans for postsecondary education to individuals who fail to register with the SSS. Nonregistrants are also ineligible for federal jobs (such as the Forest Service), federal student aid, and federal job training. A small number of colleges, especially those affiliated with the 3 “historic peace churches” (Quakers, Mennonites, Brethren), provide financial assistance to SSS nonregistrants.

WILL THERE BE ANOTHER DRAFT?
The question of whether a draft will be reinstated to supply personnel for the US military remains a matter of conjecture. There is little public support for conscription,8 and the military favors the current all-volunteer armed services. Many pundits believe that it would be impossible, politically, to reinstate such an unpopular policy. President George W. Bush and members of his administration have stated on several occasions that the current professional armed forces is effective and that they have no plans to reinstitute a draft9,10 despite the call for an imminent increase in troop strength.11 Official statements notwithstanding, administration and Pentagon officials are leaving open the possibility of a draft,10 and the secretary of Veterans Affairs recently supported a draft in public statements.12 Some form of conscription may well occur as the personnel needs of the military services grow ever more pressing in light of current and anticipated US military operations.13 Pediatricians need to be informed about these developments to counsel their patients appropriately.

IF THERE IS ANOTHER DRAFT, WHAT WILL IT LOOK LIKE?
At this point, there is only speculation as to what form a new draft would take. The most expeditious plan would be to restart the old draft system, for which an infrastructure already exists and for which young men are already cataloged in a database. Another possibility, in a less urgent setting, is that a new draft will look very different from past ones. A revised draft may be cloaked in language such as “universal” or “mandatory” national service, with military participation comprising but one option. In a new draft, several modifications may be enacted. First, the age of eligibility will likely be extended to the mid-30s or even mid-40s. Second, specific occupations will probably be targeted. At present, SSS registrants are not required to inform the agency about their profession or skills. However, with an increasing need for personnel with specialized skills (eg, persons who are fluent in certain languages, computer specialists, medical personnel including physicians), some form of targeted conscription is likely. In light of the moral ambiguities14 and lack of financial incentives involved in participating in the military, health care workers in particular have been increasingly reluctant to enlist.15 Third, women may well be included along with men, so both young women and young men need to become aware of their options.16

There have already been attempts to reinstate conscription. Representative Charles Rangel (Democrat, New York) has introduced a new draft bill during each recent session of Congress. He states that his intent is to raise awareness of the inherent injustice of the current volunteer system, which draws a preponderance of individuals from lower socioeconomic groups and persons of color. The most recent bill, House Resolution 163 (Universal National Service Act of 2003), was soundly defeated in the US House of Representatives (by a vote of 402 to 2) 1 month before the 2004 presidential election. Representative Rangel recently reintroduced the bill (House Resolution 4752) with a similar provision: that all persons 18 to 26 years old (male and female) in the United States perform 2 years of military service or other service to the country “in furtherance of the national defense and homeland security, and for other purposes.” He publicly acknowledges that the bill has little chance of passing but insists that it will raise awareness of the disparity of race and class between military recruits and the general population. A similar bill (Senate 89) languishes in the Senate, where it has been referred back to the Committee on Armed Services.

ALTERNATIVE OPTIONS TO CONSCRIPTION: CONSCIENTIOUS OBJECTION
One aspect of a new draft that has received little mention in governmental announcements is that of conscientious objection. The option to choose noncombatant or alternative service was an integral part of past drafts in our.
country. In the most recent draft, during the Vietnam War, men granted conscientious objector (CO) status were assigned to a variety of alternative public service projects within the United States, including work in mental health facilities, on public service projects, or as forest fire fighters.

According to the official SSS Web site, a CO is defined as “one who is opposed to serving in the armed forces and/or bearing arms on the grounds of moral or religious principles” (see www.sss.gov). The historic peace churches have been most prominently associated with conscientious objection, because the tenets of those religious denominations forbid killing. However, at least in past drafts, it was not necessary to belong to one of these churches or to any religious group or even to profess a religious belief, to qualify for CO status. To qualify for CO status, a person needs only to demonstrate a sincere belief in the immorality of killing in any war (not just a political objection to a certain war). Ultimately, the decision as to whether an individual met the criteria for conscientious objection was made by his local draft board.

**HOW TO ESTABLISH CO STATUS**

Choosing to become a CO is obviously a complex decision for each individual, but adolescents should be informed that this option exists, and they should be provided with information on how to establish CO status. There is currently no legal process for applying for CO status when registering with the SSS (ie, there is no “check box” on the SSS enrollment form). Nevertheless, it is strongly recommended that interested individuals compile a portfolio that documents their beliefs. Detailed information about how to identify oneself as a CO and how to gather appropriate supportive materials into a portfolio can be obtained through the Web sites of the above-mentioned peace churches (www.afsc.org/youthmil/conscientious objection, www.mcc.org/us/co, www.brethren.org/genbd/witness/ConscientiousObjection/LeadersPacket.htm) or from the Center for Conscience and War, a national clearinghouse for up-to-date information about the draft and SSS registration (www.centeron conscience.org).

In past drafts, and probably in future ones, 3 questions must be addressed (ultimately defended in written and verbal form) by an individual who is pursuing CO status. To paraphrase SSS Form 22(1): Describe the beliefs that are the basis for your claim as a CO. If appropriate, state whether those beliefs would permit you to serve in a noncombatant position in the armed forces or pay taxes for war. (2) Describe how your beliefs developed. (3) Describe how your beliefs affect the way you live your life and the type of work you do or plan to do.

Interested youth should prepare a dossier with responses to these questions and gather other relevant information, including any evidence of participation in peace-promoting activities, educational sessions, readings, or other materials that influenced their beliefs. In addition, a prospective CO should obtain letters of support from at least 3 individuals who can attest to the sincerity of the applicant’s beliefs; it is recommended that one of these letters be written by someone who does not necessarily agree with the applicant’s viewpoint but who can vouch for his or her sincerity (a pediatrician might well serve in this capacity). Interested adolescents should prepare these statements and materials now, because the time available to “prove” CO status may be as little as 10 days if a draft is enacted (see www.sss.gov).

**PROVIDING INFORMATION ABOUT THE HEALTH RISKS OF COMBAT AND OPTIONS FOR SERVICE: THE ROLE OF THE PEDIATRICIAN**

As physicians, and pediatricians in particular, we place high priority on sustaining life and improving its quality. In accordance with the life-affirming spirit of our profession, young people should be informed of alternatives to military service. Although our political and religious views span a wide spectrum, we need to acknowledge that military service is not necessarily synonymous with patriotism and that serving one’s country can be accomplished in many other ways than participation in the armed forces. Few adolescents are aware of available options, and misinformation abounds about what enlistment in the military entails (Table 1). The active presence of military recruiters in communities, schools, and homes necessitates that young people be provided a balanced viewpoint. Rule-bending and ethically questionable tactics used by some recruiters, especially with the increasing pressure to meet enlistment quotas, have been amply documented and acknowledged by the hierarchy of the military services. For example, to allay further abuses of enlistment protocols, in May 2005 the US Army temporarily suspended recruiting efforts to retain personnel in ethical recruiting practices.

As part of the No Child Left Behind Act of 2001, military recruiters have access to the names, addresses, and telephone numbers of every high school student. To continue to qualify for federal aid, schools are required to supply this information to military recruiters unless the parent specifically “opts out” by signing a special form. Schools are now required to supply this opt-out form to families under the Family Educational Rights and Privacy Act. Additional information is available at www.leavemychildalone.org.

Pediatricians can provide important information to adolescents about the potential for developing posttraumatic stress disorder (PTSD). PTSD and other mental health problems have affected 18% to 30% of Vietnam War veterans and have a current prevalence of almost
Lifelong effects of PTSD occur even without physical injury or die, the frequent chronic psychological recruit understands that he or she could incur physical injury or die, the frequent chronic psychological scars of PTSD are less well appreciated by young adults. Lifelong effects of PTSD occur even without physical trauma. Psychological consequences of PTSD can include frequent reliving of traumatic experiences, nightmares, flashbacks, avoidance of any reminders of the trauma (often leading to self-imposed isolation), anxiety symptoms, depression, substance abuse, poor concentration, inability to maintain healthy interpersonal relationships, and long-term social-adjustment difficulties such as significantly lower rates of employment, marriage, and educational attainment. It would be a disservice to adolescents who are about to enlist in the military not to make them aware of PTSD and its consequences.

CONCLUSIONS

Pediatricians and other health care professionals who are involved in the care of adolescents should familiarize themselves with and keep informed about existing laws regarding SSS registration. Such laws may change any day. Indeed, if a draft is called, we need to be aware of the details and be ready to advise adolescents appropriately. Dissemination of such information may take several forms, including personal discussions with youth, participation in local forums and workshops about career options, or developing information brochures for distribution in medical offices and clinics. Information and resources can also be obtained through the Center on Conscience and War and other Web sites as noted above. A new draft might contain provisions markedly different from previous regulations, but in the meantime, it is our responsibility to provide our young patients with the opportunity to choose to follow their conscience in whichever direction it leads them.

REFERENCES

DENIAL MANAGEMENT: FIGHTS OVER HEALTH CLAIMS SPAWN A NEW ARMS RACE—INSURERS AND DOCTORS ARE SPENDING BILLIONS

“Boston-based Athenahealth Inc [is] one of the biggest of hundreds of companies in a lucrative niche: helping doctors wring payments from health plans. Athenahealth’s software flagged and corrected the complex coding for thousands of claims, preventing them from getting hung up in insurers’ Byzantine rules. . . . ‘The insurers out-code us, they out-smart us and they have more manpower,’ says Shari Reynolds, the administrator at Paluxy Valley, which pays Athenahealth a little over 3% of the $2.5 million it collects annually from insurers. ‘Now at least we have a fighting chance.’ . . . Doctors increasingly complain that the insurance industry uses complex, opaque claims systems to confound their efforts to get paid fairly for their work. Insurers say their systems are designed to counter unnecessary charges and help keep down soaring health-care costs. Like many tug-of-wars over the health-care money pot, the tension has spawned a booming industry of intermediaries. . . . It’s called ‘denial management.’ Doctors, clinics and hospitals are investing in software systems costing them each hundreds of thousands of dollars to help them navigate insurers’ systems and head off denials. . . . The battle is costing medical providers and insurers around $20 billion—about $10 billion for each side—in unnecessary administrative expenses, according to a 2004 report by the Center for Information Leadership, a non-profit health technology research group based in Boston. Some companies are profiting from arming both sides.”

Fuhrmans V. Wall Street Journal. February 14, 2007
Noted by JFL, MD
Atherosclerosis first appears during childhood. Therefore, interventions may need to be started much earlier, and targeted at biological pathways that determine disease susceptibility, to prevent adult cardiovascular disease. In this issue of Pediatrics, Martin et al put forward the hypothesis that folic acid intake during in utero and early postnatal life may be relevant to longer-term risk of cardiovascular disease. This suggestion is based on their observation that folate levels in the immediate postnatal period are associated with endothelial function.

The endothelium has a key biological influence on the development of cardiovascular disease. Endothelial responses relate to an individual’s risk of cardiovascular events, and in animal models, loss of endothelial function leads to accelerated atherosclerosis and the development of hypertension and glucose intolerance. Approximately 10 years ago it was observed that there was significant variation in endothelial function between individuals during childhood. This variation did not relate to levels of classical cardiovascular risk factors, an important determinant of function in adults and those at high risk, but did relate to patterns of early growth. Reduced in utero growth was associated with impaired endothelial function, and this relationship has been consistently demonstrated in cohorts aged from a few days of life to early adulthood. Interestingly, this observation is strikingly similar to the associations between reduced in utero growth, increased mortality from atherosclerotic cardiovascular disease, and the development of risk factors such as hypertension and glucose intolerance derived from large population studies.

These findings have led to the hypothesis that early growth and endothelial function are biologically interrelated during early life and that “programmed” variation in endothelial function determines long-term risk of cardiovascular disease (Fig 1). In support of this hypothesis, animal models have demonstrated that environmental and genetic influences can determine both early growth and endothelial function.

**MODIFICATION OF ENDOTHELIAL FUNCTION IN EARLY LIFE**

If early variation in endothelial function does determine risk of later cardiovascular disease, then the endothelium is a novel target for pediatric prevention strategies. There has been particular interest in folic acid as an intervention. It is a simple nutritional supplement that influences endothelial function by having either an effect on homocysteine metabolism or a more direct effect on endothelial nitric oxide synthase function.

Interestingly, low folic acid levels in the mother are also associated with a reduced birth weight in some studies, and animal data have demonstrated that vascular dysfunction in infants (secondary to maternal malnutrition) can be prevented by supplementation with folic acid during pregnancy.

Martin et al have demonstrated, for the first time in humans, that folate levels of both the mother and infant relate to endothelial function in the infant during the early postnatal period. Furthermore, they demonstrated that low folate levels in the mother, but not the infant, are related to reduced birth weight. However, they did...
not demonstrate that variation in folate levels accounts for the relation between birth weight and endothelial function, and it remains probable that there are other important modifiable pathways.

REMAINING ISSUES

Martin et al have provided data to focus attention on folic acid biology as a potentially simple target for pediatric prevention of atherosclerosis. However, at present, advice for folic acid supplements during pregnancy should be based on their known benefits for neurologic development and not on any hypothesized long-term cardiovascular protection to the offspring, because there remain at least 2 key issues. First, there are no data available to confirm that variation in endothelial function during the first decades of life has a long-term impact on structural atherosclerosis or emergence of risk factors in humans. Second, the influence of folic acid on endothelial function is known to be transient, with acute changes in levels associated with acute changes in function. This contrasts with the relations between in utero growth, endothelial function, and cardiovascular disease, which are fixed over decades.7,12 Therefore, we still need to unravel the underlying biology to determine if transient changes in folic acid intake can lead to long-term “programmed” variation in endothelial function and cardiovascular risk.

Over the next few years, these issues will be addressed by current longitudinal studies of the impact of early endothelial function on the vasculature combined with basic research into the effect of interventions on endothelial function and disease susceptibility. Together, a growing body of research into the effect of early variation in vascular biology on long-term risk of cardiovascular disease will generate a scientific evidence base to allow us to design robust proposals for novel pediatric prevention strategies for atherosclerosis.

REFERENCES


Providing a Safety Net for Children: Investigating a Multistate Outbreak of *Ralstonia mannitolilytica* Related to a Contaminated Reusable Device

Lisa Saiman, MD, MPH

Department of Pediatrics, Columbia University, New York, New York; Department of Epidemiology, Morgan Stanley Children’s Hospital of New York-Presbyterian, New York, New York

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

*Jhung et al*¹ have reported a multistate outbreak of *Ralstonia mannitolilytica* that contaminated a Vapotherm 2000i, a reusable oxygen-warming device. As of 2005, the Vapotherm 2000i was used in 900 American hospitals, primarily in NICUs.¹ In the outbreak described in their report, 38 children (5 days to 7 years of age) were infected or colonized with *R. mannitolilytica*; fortunately, there was only 1 death attributed to this very unusual pathogen. In the case-control study performed (4 controls per case-subject matched on length of hospital stay), the case-subjects were 18-fold more likely than control subjects to have been exposed to a Vapotherm 2000i.¹ To confirm the association of this outbreak with the Vapotherm 2000i, molecular analysis of isolates from 18 hospitals in 12 states was performed by investigators from the University of Michigan. These studies (1) confirmed the identity of this unusual species, (2) determined that putative case-subjects were infected/colonized with the same clone of *R. mannitolilytica*, (3) determined that the clone recovered from the patients was also recovered from the Vapotherm 2000i, and (4) documented the genetic diversity of *R. mannitolilytica*. These epidemiologic and molecular data were the basis of the product recall (the Vapotherm 2000i was recalled from use in December 2005).

Consider the magnitude of the investigation and the implications of recalling a product in widespread use in hospitals. The magnitude of the investigation is reflected by the lengthy list of coauthors and acknowledgments in the Jhung et al report. Coauthors included members of the Centers for Disease Control and Prevention (CDC) Division of Healthcare Quality Promotion, infection-control practitioners from 2 children’s hospitals, and epidemiologists from the Philadelphia health department, and an additional 22 children’s hospitals and health departments were listed in their acknowledgments. The list is open testimony to the commitment of these institutions to improve the safety of health care for children.

It is fair to say, generally, that outbreak investigations, case-control studies performed to assess potential risk factors for infection, and molecular epidemiology studies to determine clonality are found within the pages of infection-control subspecialty journals. It is also fair to say that descriptions of emerging and/or unusual pathogens are found within the pages of infectious-diseases subspecialty journals. So why publish the description of an outbreak involving an unusual pathogen in a limited number of children and the measures taken to confirm the role of a reusable device in the journal *Pediatrics*? This publication educates the pediatric community about the safety net cast by the infection-control community to protect children from health care–associated (formerly termed nosocomial) infections. Furthermore, this publication educates the pediatric community about the complex interactions between children’s health care facilities, the CDC, the Food and Drug Administration (FDA), and industry.

Opinions expressed in these commentaries are those of the authors and not necessarily those of the American Academy of Pediatrics or its Committees.

References:


Address correspondence to Lisa Saiman, MD, MPH, 622 W 168th St, PH 4 West, Room 470, New York, NY 10032. E-mail: ls5@columbia.edu

PEDIATRICS (ISSN Numbers: Print, 0031–4005; Online, 1098–4275). Copyright © 2007 by the American Academy of Pediatrics
Jhung et al have presented not only the complexities of outbreak investigations but also ably demonstrated the multiple levels of communication and collaboration required to ensure a productive investigation. Throughout the United States, infection-control practitioners and clinicians who care for high-risk pediatric patients perform surveillance for health care–acquired infections and colonization. Colonization can be as significant as infection; colonized children can serve as reservoirs of potential pathogens for other vulnerable patients. Thanks to continual surveillance in individual institutions, infection-control practitioners acquire a tremendous amount of pattern recognition. One child infected (or colonized) with an unusual pathogen signals a fluctuation in the seismographic recording; 2 cases trigger a reading akin to a tremor and often prompt a realignment of priorities in case there is an earthquake brewing; and 3 such children require activation of an “emergency response team.”

As illustrated by their report, health care facilities communicate epidemiologically significant events to their local health departments, who in turn invite the CDC to assist in outbreak investigations, particularly when a multistate outbreak is suspected. Multiple examples of such investigations have been performed by the CDC in efforts to document the links between specific exposures to potential pathogens and illness. Recent examples include a multistate outbreak of *Enterobacter sakazakii* linked to contaminated infant formula, a multistate outbreak of *Pseudomonas fluorescens* linked to contaminated normal saline syringes used to flush indwelling central venous catheters, and a multistate investigation that explored the increased rate of bloodstream infections, particularly with Gram-negative bacilli, associated with continuous intravenous therapy for pulmonary hypertension. Each of these examples identified new risk factors and serve to illustrate the complexities of ensuring safe health care for increasingly vulnerable patients who are increasingly dependent on technology.

Although the manufacturing and disinfecting protocols for the Vapotherm 2000i are somewhat intricate, it is worthwhile to ponder them. All health care institutions devote considerable resources to ensuring appropriate levels of disinfection and sterilization. Patient care items and equipment are divided into 3 categories (critical, semicritical, and noncritical) that reflect the degree to which items and equipment confer a risk of infection. Critical items and equipment require sterilization, because such items come into contact with sterile tissues or the vascular system and any level of contamination confers a risk of infection with spores, mycobacteria, fungi, viruses, and/or bacteria. Examples of such items are surgical instruments, urinary tract catheters, and vascular catheters. Semicritical items contact mucous membranes or nonintact skin and require high-level disinfection with processing that kills mycobacteria, fungi, viruses, and bacteria. Examples of such items include respiratory therapy equipment, bronchoscopes, and endoscopy equipment. Noncritical items (eg, blood pressure cuffs and crutches) only come into contact with intact skin and, thus, are associated with a minimal risk of infection to the patient and require low-level disinfection. Thus, the Vapotherm 2000i is a semicritical medical device, and its reusable components must undergo high-level disinfection between patients. However, the manufacturer’s instructions mistakenly specified low-level disinfection between patients. In addition, the filter that separates air to the patient from the water used to humidify the air was intended to serve as a total biological barrier; thus, tap water was used to fill the chamber. Tap water is not sterile and has been linked to infections with *Legionella*, mycobacteria, enteric pathogens, noroviruses, and cryptosporidium. Many of these potential pathogens form biofilms within the water pipes and tap that cannot be eradicated. The investigators never identified the “smoking gun” in that *Ralstonia* was not recovered from the tap water at the manufacturing plant, but it is notable that these cultures were performed months after the contaminated equipment was manufactured, and other potential pathogens (*Sphingomonas paucimobilis* and *Burkholderia cepacia*) did grow from the tap water.

Around the world, the health care community is struggling with the safety of reusable devices that are intended to help reduce health care costs and reduce waste. Although the tremendous cost of single-use items for patients is duly noted, the potential risks associated with reusable devices are not fully quantified. Furthermore, many institutions and professional societies are also struggling with the safe reprocessing of single-use devices. Without this publication and others like it, the pediatric community would be unable to monitor and measure the risk benefits of reusable patient care items and equipment.

Finally, the Jhung et al report disclosed another significant factor that contributed to this outbreak: the mechanism whereby the FDA clears new medical devices for marketing. The FDA has a premarket approval process that requires demonstration of safety and effectiveness of certain new or high-risk devices before approval for marketing. In contrast, the FDA has a 501(k) process whereby a new medical device can be marketed if deemed to perform equivalently to an approved, currently marketed device and has the same intended use. According to the FDA’s *Report on New Medical Devices Approved in Fiscal Year 2000*, the vast majority of medical devices (nearly 99%) were cleared for usage by the 501(k) process. The Jhung et al investigation highlights the potential noncomparability of new devices as a result of reprocessing issues. Thus, the FDA, in recognition of the importance of infection-control principles, has made...
device sterilization a priority for improving the 501(k) submission process.8

In summary, Jhung et al carefully outlined the productive collaboration between children’s health care institutions, governmental agencies, and industry. Their report traced the complex steps taken from the clinical observations of astute infection-control practitioners to reporting potential outbreaks to local health departments to engaging the CDC and conferring with the FDA. Also, as their report detailed, progress can only be made when the manufacturer is involved as a participant in improving health care. Finally, each institution must have a careful process whereby all new patient care items and equipment are brought into usage with consideration for the infection-control implications.

REFERENCES
Maternal Smoking, Asthma, and Bronchiolitis: Clear-Cut Association or Equivocal Evidence?

Eric A. F. Simões, MB, BS, DCH, MD
Department of Pediatrics, University of Colorado at Denver and Health Sciences Center and Children's Hospital, Denver, Colorado

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

Although there is no doubt that secondhand environmental tobacco smoke exposure contributes to lower respiratory tract infection (LRTI) in infants and young children in developing countries and industrialized nations, the evidence for a specific effect on respiratory syncytial virus (RSV) LRTI is less clear. It is the equivocal nature of the evidence for the role of smoking affecting the severity of RSV bronchiolitis that has, in part, prevented the American Academy of Pediatrics from strongly advocating smoking as a risk factor when considering prophylaxis with palivizumab for previously premature infants of 32 to 35 weeks' gestation.

In this issue of Pediatrics, Carroll et al present probably the largest population-based study of term infants with bronchiolitis to determine the association between maternal asthma, maternal smoking, and the incidence and severity of bronchiolitis. In their cohort of > 100,000 mother-infant dyads, > 20% of the infants had 1 health care visit for bronchiolitis. Maternal smoking increased the risk of bronchiolitis by 14%, maternal asthma raised it by 39%, and both together raised it by 47%. Infants of mothers with asthma who smoked had a higher risk for emergency department visits and hospitalizations. Although maternal smoking increased the risk of prolonged hospitalization by 19%, if the mother was smoking and had asthma the increased risk was 38%. Given the enormous size of the study encompassing the vast majority of children, infants, and mothers enrolled in the Tennessee Medicaid program between 1995 and 2003, this would seem to be compelling evidence for an association between maternal smoking and bronchiolitis.

Despite the enormous size of this study, the results must be tempered with an understanding of the study limitations, which were acknowledged by the authors.

The major limitation is that other risk factors that are well known to be associated with hospitalization for RSV or bronchiolitis—such as an index of crowding that includes the number of sibling and adults in the household, day-care exposure, or day-care exposure of siblings or the presence of siblings < 5 years of age, birth during the first half of the RSV season, and a putative risk factor (breastfeeding)—were not included in these analyses.

In an examination of atopic risk factors, although a maternal history of asthma was considered, that of other relatives or other conditions such as atopic dermatitis or allergic rhinitis in first-degree relatives was not included in the analysis. Finally, a diagnosis of bronchiolitis does not equate to RSV, although in RSV season the majority of bronchiolitis admissions (perhaps 60%–80%) would be caused by RSV.

An initial study in Rochester, New York, conducted between 1974 and 1976, suggested that there is a significant difference between children hospitalized with RSV and controls in the amount of environmental tobacco smoke exposure (76% vs 40%, respectively; P < .01). In that study, unfortunately, no multivariate analysis was performed. A subsequent case-control study of 53 subjects with bronchiolitis and 106 matched controls showed an increased risk of bronchiolitis in families with ≥ 1 smoker. In this study, a multivariate analysis that included a socioeconomic index, breastfeeding, crowd-

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.2007-0869
doi:10.1542/peds.2007-0869
Accepted for publication Mar 19, 2007
Address correspondence to Eric A. F. Simões, MB, BS, DCH, MD, Children’s Hospital, 1056 E 19th Ave, B070, Denver, CO 80218. E-mail: eric.simoes@uchsc.edu
ing, number of siblings in the home, and a family history of asthma showed a significant effect of passive smoke exposure in those with or without a family history of asthma. Several other prospective case-control studies from Sweden, Finland, and Oakland, California, showed no effect, but they were designed to examine risk factors for the development of subsequent asthma and did not analyze many risk factors for RSV hospitalization. In a large case-control study from Denmark that was designed to examine risk factors for RSV hospitalization, smoking during pregnancy was one of the significant factors on multivariate analysis (odds ratio: 1.56; 95% confidence interval: 1.32–1.98). Recent studies from Alaska and Australia showed a significant effect on univariate analysis (P = .018 and .0004, respectively) but did not show the same effect on multivariate analysis. In the study from Alaska, risk factors included breastfeeding, crowding, siblings in the home, and day-care exposure but did not include siblings in day care or maternal history of asthma or other family members with atopy. In contrast, the case-control study from Australia did not examine any of the other well-known risk factors but examined risk factors not commonly known to be associated with RSV hospitalization. Not surprisingly, the multivariate analysis showed no effect of smoking in this population.

Should one then look at prospective cohort studies to provide this answer? Clearly the most important was the one conducted in Tucson, Arizona. In this large study of 1179 infants that examined risk factors for the development of RSV LRTI, a multivariate analysis showed no significant effect of environmental tobacco smoke exposure. In a more recent nested case-control study from the Danish birth cohort that involved 2564 infants and children who were hospitalized with RSV and 12 816 age-matched controls (followed from birth to 18 months of age), tobacco smoke exposure was shown to be unequivocally associated with an increased risk of hospitalization with RSV (odds ratio: 1.35; 95% confidence interval: 1.20–1.52). In this study, in which individual-level patient data were obtained and every one of the known and suspected risk factors described above was included, environmental tobacco smoke exposure was shown to increase the risk of hospitalization in infants in each of the age groups examined: 0 to 5, 6 to 11, and 12 to 18 months of age. This study also established the role of an atopic disposition for hospitalization of infants with RSV bronchiolitis: the adjusted relative risk of RSV hospitalization in the offspring was 1.11 for maternal atopic dermatitis, 1.72 for maternal asthma, and 1.23 for paternal asthma.

In none of the above-mentioned case-control or cohort studies, or from the study by Carroll et al., was the major question answered: Does environmental tobacco smoke exposure in prematurely born infants increase their risk for RSV hospitalization? In fact, the Carroll et al. study excluded such children and focused on “normal” children. Two prospective cohort studies from Spain (in infants <32 weeks’ gestation), 19,20 1 study from Germany (in infants <35 weeks’ gestation), 21 1 study from Canada (in infants 33–35 weeks’ gestation), 22 and a case-control study from Spain (in infants 33–35 weeks’ gestation) have addressed this issue in premature infants. All of these recent studies included most of the known risk factors for RSV hospitalization. Only 2 of these 5 studies, both from Spain, 19,22 showed an independent risk for environmental tobacco smoke exposure on RSV hospitalizations. The article by Carroll et al. and both the large Danish studies 9,15 fairly convincingly support the role of environmental tobacco smoke exposure as an important risk factor for development of severe RSV disease leading to hospitalization. However, neither of these studies addressed the critical question to which pediatricians would like to know the answer: Is the premature infant at such a high risk of RSV hospitalization from prematurity alone, and does tobacco smoke exposure increase this risk further?

REFERENCES


“In Medicine one must pay attention not to plausible theorizing but to experience and reason together. . . .
I agree that theorizing is to be approved, provided that it is based on facts, and systematically makes its deductions from what is observed. . . . But conclusions drawn from unaided reason can hardly be serviceable; only those drawn from observed fact.”
—Hippocrates, Precepts

The Success of an Immunization Information System in the Wake of Hurricane Katrina

Julie A. Boom, MD, Anna C. Dragsbaek, JD, Cynthia S. Nelson, MPH

ABSTRACT
Within days after Hurricane Katrina in September 2005, the Houston-Harris County Immunization Registry was connected to the Louisiana Immunization Network for Kids Statewide. This linkage provided immediate access to the immunization records of children who were forced to evacuate the New Orleans, Louisiana, area. One year later, >18 900 immunization records have been found, representing an estimated cost savings of more than $1.6 million for vaccine alone and $3.04 million for vaccine plus administration fees. This experience demonstrated the vital and previously unrecognized functionality of immunization information systems in a public health emergency. Here we describe the Houston-Harris County Immunization Registry’s experience after Hurricane Katrina in terms of maximizing the use of immunization information systems and the implications of this experience for patients, providers, and public health for future disaster-preparedness planning.

IMMUNIZATION INFORMATION SYSTEMS (IISs) are confidential, population-based computerized information systems that collect immunization data for all children and, in some cases, adults within a geographic area. IISs are currently under development or are actively being used in every state in the United States. An innovative use of IISs was achieved during the public health emergency after Hurricane Katrina.

Before Hurricane Katrina, the development and use of IISs centered on several key purposes:

1. consolidating children’s immunization histories that have been fragmented between multiple providers;
2. helping providers to “forecast” which immunizations are due for a child;
3. recalling children who are due or overdue for immunizations;
4. assisting providers to determine immunization coverage levels;
5. generating complete and official immunization certificates;
6. identifying pockets of need in the community; and
7. avoiding underimmunization or overimmunization.

Participation in IISs has steadily increased over the past decade. In 1999, 46% of public providers and 13% of private providers used IISs compared with 75% of public providers and 44% of private providers in 2005 (D. Bartlett, MPH, Centers for Disease Control and Prevention [CDC], written communication, October 10, 2006). The Healthy People 2010 goal for IIS participa-
tion is ≥95% of children younger than 6 years. The number of children younger than 6 years with at least 2 doses recorded in an IIS has increased from 32% in 1999 to 56% in 2005 (D. Bartlett, MPH, CDC, written communication, October 10, 2006).

IISs have been found to be most useful in consolidating fragmented records from different providers in the community, region, or state. One study suggests that 25% of children receive immunizations from more than 1 provider by the age of 3 years. As a result, providers often have incomplete or inaccurate immunization records. Such fragmentation of immunization records leads to missed opportunities, duplicate immunization, and inaccurate measurement of coverage levels. In 1 study, researchers used registry immunization records from private-provider practices, community health centers, and public health clinics to determine the impact of a regional registry on overall up-to-date rates for children aged 24 months. As more children aged 7 to 24 months visited multiple provider sectors, they found that record fragmentation among the 3 sectors increased. When they sequentially added immunization-registry records from community health centers and public health clinics to records of children who had immunizations recorded at private practices, the relative increase in the overall up-to-date rate ranged from 9% for children 7 months of age to 50% for children 24 months of age.

In addition, IISs assist with the management of vaccine supply, reduce missed opportunities, and generate complete vaccine records for patients, parents, providers, and schools. By using forecasting algorithms, IISs can also identify which immunizations are due. This basic function facilitates the implementation of reminder/recall systems, which have been shown to improve immunization rates by 1% to 20% in a variety of medical settings.

By decreasing immunization record fragmentation and using functionalities of IISs such as forecasting and reminder/recall, providers can eliminate underimmunization or overimmunization of children. Feikema et al found that 21.1% of children were overimmunized for at least 1 vaccine in 1997, which represented ~1.8 million extra doses. Results from an assessment of the benefits from using an immunization registry to determine inner-city immunization rates for infants showed that 19% of the study infants had received additional immunizations. Reducing the number of duplicate immunizations administered annually by using state- and community-based IISs would result in a cost savings of $26.5 million for extra doses administered, vaccine waste, and extra clinic visits.

**NEWLY RECOGNIZED VALUE OF IISs**

A new and powerful utility for IISs emerged in the wake of Hurricane Katrina. In the immediate days after this devastating event, ~200 000 evacuees from the greater New Orleans area headed for Houston, Texas. Most evacuees arrived with few personal belongings. Needless to say, most children’s immunizations records were left at home.

In an effort to assist patients’ families to find their children’s records, the staff of the Houston-Harris County Immunization Registry (HHCIR) contacted the vendor for HHCIR and the Louisiana Immunization Network for Kids Statewide (LINKS) to inquire about connecting the 2 systems. Importantly, both registries were already following Health Level Seven (HL7) standards. HL7 is a nationally recognized standard for electronic exchange of health-related data between computer systems. Using HL7, a bridge was built between the 2 IISs in <24 hours. Programmers diligently tested the connection, and the HHCIR-LINKS connection went live on September 9, 2005, a mere 10 days after landfall of the hurricane.

To assist public health officials in searching for children’s immunization records, the Web-based immunization registry was made available to providers in the makeshift medical clinics in the Astrodome and George R. Brown Convention Center. Health care workers at these shelters were able to readily access records for children by using the HHCIR-LINKS connection. Figure 1 shows the cumulative number of LINKS immunization records searched and found according to month using

![Figure 1](https://example.com/figure1.png)

**Cumulative number of LINKS immunization records searched and found in the months after Hurricane Katrina (September 2005 to September 2006).**
the HHCIR-LINKS connection. One year later, 18,966 records have been found for Louisiana residents, both children and adults, because LINKS is a lifetime IIS.

The HHCIR-LINKS connection prevented children from receiving duplicate immunizations. Immediately after the disaster, children were allowed to enroll provisionally in Texas schools without proof of immunization. This provisional enrollment expired on October 31, 2005, approximately 9 weeks after the disaster. Figure 2 shows the number of LINKS immunization records found through the HHCIR-LINKS connection each month after Hurricane Katrina. As Fig 2 demonstrates, this HHCIR-LINKS connection continued to be used after the expiration of provisional enrollment, with an increase at the beginning of the spring semester in January 2006 and again at the beginning of the new school year in August 2006.

**SAVINGS ASSOCIATED WITH AVOIDING OVERIMMUNIZATION**

Every LINKS record recovered through the HHCIR-LINKS connection represents money saved from revaccination of these children. If immunizations were re-administered to every child for whom a record was found, the cost for vaccine supply alone would total more than $1.60 million (Table 1). In Table 1, the Vaccines for Children Program price list (current as of October 31, 2006) was applied for all immunizations recovered through this connection. As the table demonstrates, if a vaccine administration fee of $14.85 per dose were assumed (amount permitted per the Vaccines for Children Program), the cost of vaccine plus administration fees would total $3.04 million. These figures are based solely on the immunization histories included in the records found through the Houston connection. LINKS data were available to providers nationwide. National cost data were analyzed by Urquhart to calculate the savings incurred by all locations that accessed LINKS data (G. Urquhart, MPH, CDC, verbal communication, November 6, 2006).

The costs calculated here reflect the costs associated with avoiding overimmunization, which denotes only a partial cost analysis for the connection created between HHCIR and LINKS. Although a net cost analysis is beyond the scope of this report, these calculations demonstrate that IISs can avoid some of the significant costs associated with overimmunization, a likely problem in the event of a natural disaster in which medical charts are destroyed. By using an IIS in these situations other costs savings would be realized. For example, vaccine shipment, storage, and handling, hiring of additional personnel, and allocation of physical space to administer shots to a large number of children would be avoided. Alternatively, costs were incurred to implement the HHCIR-LINKS connection, such as technical fees and training of additional personnel to use the IIS. Furthermore, the cost savings we report are a high-range estimate. It is likely that not all children for whom records were found would need all immunizations repeated, because some vaccines are not administered to older children and adults. However, this estimate of cost savings demonstrates the potential economic benefit of using IISs in public health emergencies.

**IMPLICATIONS**

The IIS experience after Hurricane Katrina demonstrated an important use for IISs: they can serve as an important safety net for patients, providers, and the community by storing immunization records that might otherwise be lost or destroyed. Using an IIS in such situations can help to avoid the costs associated with revaccinating children who have lost their immunization records.

**Benefits to Patients**

From a patient perspective, an IIS is a valuable tool to safeguard important health information that will be used throughout a person’s life. Paper records that are stored in homes may be lost because of natural disasters, such as the flooding that occurred in Hurricane Katrina, or destroyed by fire, earthquakes, or other natural and manmade disasters. Furthermore, using IISs in these situations saves thousands of children from the pain and inconvenience of reimmunization, the cost of which does not have a tangible price tag. In addition, the HHCIR-LINKS connection allowed parents to avoid other

**FIGURE 2**

Number of LINKS immunization records found according to month after Hurricane Katrina (September 2005 to September 2006).
frustrations and costs associated with getting their children reimmunized, such as lost work time and missed school.

Benefits to the Provider
IISs serve as an important safety net to providers by preserving the medical chart. Providers in both Texas and Louisiana realized the importance of such a safety net after Hurricane Katrina. After the hurricane, >18,000 immunization records for Louisiana children and adults were found through the HHCIR-LINKS connection. Recovery of these records saved providers in Houston and Louisiana the staff time and vaccine costs that would have been required to reimmunize these children. However, there were many children whose immunization records were not available through the HHCIR-LINKS connection. As a result, providers assisted families with recovery of immunization records by contacting their Louisiana medical homes or schools, which strained limited staff time and resources. In many cases the records were permanently lost, and providers were forced to reimmunize children. By preserving the medical chart with an IIS, providers are able to serve efficiently and effectively the needs of their patients when paper records are permanently lost.

Benefits to Public Health
Finally, IISs provide an invaluable public health tool to our community. Using IISs, we can avoid revaccinating large numbers of children in the wake of a catastrophic loss of medical charts. Because newer vaccines are increasingly expensive and more vaccines are required for school entry, the loss of immunization records can represent an enormous expense to our private and public health systems. The data recovered through the HHCIR-LINKS connection alone represented a huge cost savings to the community.

Hurricane Katrina demonstrated the usefulness of an IIS in a public health crisis involving loss of medical charts. To realize the full value of these systems in disaster situations, certain conditions must be met. First, IISs can only serve the role of a safety net if they are populated and used by patients and providers. Second, they must be technically cohesive if they are to be linked as occurred during Hurricane Katrina. Public health officials need to work toward all IISs being HL7 compliant and following the functional standards as recommended by the American Immunization Registry Association and the CDC.

Some IISs, such as the Michigan Care Improvement Registry (MCIR), have already added several functionalities that could be used in the event of a public health emergency. For example, the MCIR has an all-hazard module to quickly capture immunization data during a mass immunization campaign should an outbreak occur. Included in the all-hazard module is a “recipient tier grouping” that allows public health officials to record immunizations and medication administration data on first responders, health care workers, and essential workers during an epidemic such as pandemic influenza.

CONCLUSIONS
Our experience after Hurricane Katrina can be generalized to include any natural disaster or public health emergency. For example, public health officials can use registry data to identify pockets of need to prevent outbreaks and manage those areas should an outbreak occur. In addition, IIS functionalities such as those implemented by the MCIR facilitate effective and efficient preparation for and response to various types of public health emergencies.

<table>
<thead>
<tr>
<th>Vaccine</th>
<th>N</th>
<th>Vaccine Price, $</th>
<th>Total With No Administration Fee, $</th>
<th>Total With $14.85 Administration Fee, $</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTaP</td>
<td>22,638</td>
<td>12.25</td>
<td>277,316</td>
<td>613,490</td>
</tr>
<tr>
<td>DTaP-Hep B-IPV</td>
<td>544</td>
<td>43.75</td>
<td>23,800</td>
<td>31,878</td>
</tr>
<tr>
<td>DTaP-Hib</td>
<td>621</td>
<td>24.94</td>
<td>15,488</td>
<td>24,710</td>
</tr>
<tr>
<td>Hep A</td>
<td>131</td>
<td>12.10</td>
<td>1,240,088</td>
<td>326,582</td>
</tr>
<tr>
<td>Hep B</td>
<td>13,636</td>
<td>9.10</td>
<td>218,000</td>
<td>312,657</td>
</tr>
<tr>
<td>Hib</td>
<td>13,690</td>
<td>7.92</td>
<td>118,000</td>
<td>241,012</td>
</tr>
<tr>
<td>Hib-Hep B</td>
<td>2292</td>
<td>26.25</td>
<td>60,165</td>
<td>94,201</td>
</tr>
<tr>
<td>M, R, MMR, Mu, M/R</td>
<td>9731</td>
<td>17.28</td>
<td>168,152</td>
<td>312,657</td>
</tr>
<tr>
<td>Varicella</td>
<td>18</td>
<td>68.00</td>
<td>12,244</td>
<td>1491</td>
</tr>
<tr>
<td>Pneumococcal</td>
<td>5551</td>
<td>57.59</td>
<td>319,682</td>
<td>402,114</td>
</tr>
<tr>
<td>Polio</td>
<td>19,841</td>
<td>10.82</td>
<td>214,680</td>
<td>509,318</td>
</tr>
<tr>
<td>DT, Td</td>
<td>2131</td>
<td>16.62</td>
<td>35,417</td>
<td>67,063</td>
</tr>
<tr>
<td>Varicella</td>
<td>4374</td>
<td>56.90</td>
<td>248,881</td>
<td>313,835</td>
</tr>
<tr>
<td>Total</td>
<td>96,482</td>
<td>—</td>
<td>1,609,197</td>
<td>3,042,192</td>
</tr>
</tbody>
</table>

DTaP indicates diphtheria-tetanus-acellular pertussis; Hep A, hepatitis A; Hep B, hepatitis B; IPV, inactivated poliovirus vaccine; Hib, Haemophilus influenzae type b; M, measles; R, rubella; MMR, measles-mumps-rubella; Mu, mumps; DT, diphtheria and tetanus toxoids; Td, tetanus and diptheria.
health emergencies. Although these capabilities are not needed on a daily basis, we live in a world in which public health preparedness is vital. When IISs are fully used in a public health emergency, they yield economic benefits and strengthen our ability to serve our patients when they are most vulnerable.

ACKNOWLEDGMENT
We thank Dr Alan Hinman for his time and effort in reviewing this manuscript.

REFERENCES
POLICY STATEMENT

Provision of Educationally Related Services for Children and Adolescents With Chronic Diseases and Disabling Conditions

Council on Children With Disabilities

ABSTRACT
Children and adolescents with chronic diseases and disabling conditions often need educationally related services. As medical home providers, physicians and other health care professionals can assist children, adolescents, and their families with the complex federal, state, and local laws, regulations, and systems associated with these services. Expanded roles for physicians and other health care professionals in individualized family service plan, individualized education plan, and Section 504 plan development and implementation are recommended. Recent updates to the Individuals With Disabilities Education Act will also affect these services. Funding for these services by private and nonprivate sources also continue to affect the availability of these educationally related services.

The complex range of federal, state, and local laws, regulations, and systems for special education and related services for children and adolescents in public schools is beyond the scope of this statement. Readers are referred to the American Academy of Pediatrics policy statement "The Pediatrician's Role in Development and Implementation of an Individual Education Plan (IEP) and/or an Individual Family Service Plan (IFSP)" for additional background materials. The focus of this statement is the role that health care professionals have in determining and managing educationally related services in the school setting.

This policy statement is a revision of a previous statement, “Provision of Educationally Related Services for Children and Adolescents With Chronic Diseases and Disabling Conditions,” published in February 2000 by the Committee on Children With Disabilities (http://aappolicy.aappublications.org/cgi/content/full/pediatrics;105/2/448).

FEDERAL LAWS
Related services such as speech therapy, occupational therapy, physical therapy, and nursing care are provided to students in school because they are related to the student’s education. The term “related services” as currently defined in Part A of the Individuals With Disabilities Education Act (IDEA) includes the following:

. . . transportation and such developmental, corrective, and other supportive services (including speech-language pathology and audiology services, psychological services, physical and occupational therapy, recreation, including therapeutic recreation, social work services, counseling services, including rehabilitation counseling, orientation and

www.pediatrics.org/cgi/doi/10.1542/peds.2007-0885
doi:10.1542/peds.2007-0885
All policy statements from the American Academy of Pediatrics automatically expire 5 years after publication unless reaffirmed, revised, or retired at or before that time.

Key Words
IEP, IDEA, Section 504, related services, special education, ICDH-2, physician’s role, children with chronic diseases and disabilities

Abbreviations
IDEA— Individuals With Disabilities Education Act
IEP— Individualized education plan
PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275). Copyright © 2007 by the American Academy of Pediatrics
mobility services, and medical services, except that such medical services shall be for diagnostic and evaluation purposes only) as may be required to assist a child with a disability to benefit from special education, and includes the early identification and assessment of disabling conditions in children.

The legal justification for the provision of related services without qualifying for special education placement can be found in Section 504 of the Rehabilitation Act of 1973. This section prohibits discrimination that is based on disability within federal and federally assisted programs. Regulations promulgated by the US Department of Education have more broadly defined the persons covered by this act, as well as the services that are to be provided. According to Section 504, all children must be provided with an appropriate education that “could consist of education in regular classes, education in regular classes with the use of supplementary services, or special educational and related services.” Psychological testing and evaluation, counseling, physical and occupational therapy, medical services, speech pathology, audiology, and orientation mobility instruction are listed among the types of “developmental, corrective, and . . . support services” that may be provided to qualified persons. Thus, Section 504 states that children with chronic diseases and disabling conditions are entitled to appropriate modifications within their educational program to accommodate their special needs regardless of whether their classroom placement is considered regular education or special education. Some school systems have developed flexible, function-oriented “504 modification plans” for students. Unfortunately, some school systems still provide few services. Children with chronic medical conditions, who usually function well in the standard classroom, still need consideration for related services. Examples of such children are those with asthma and allergies, who often find themselves at odds with their schools and school districts because of issues related to classroom modifications (eg, no pets in the classroom, having hand-washing facilities), curriculum modifications (eg, alternatives to standard physical education on an as-needed basis rather than the usual exclusion or full participation), and access to medications.

On December 3, 2004, the IDEA (Pub L No. 108–446) was enacted. Most of the provisions of this law became effective July 31, 2005. The new law is likely to have a major impact on how students with disabilities are educated. Listed below are some of the key changes that occurred with the IDEA 2004.

- The long-established obligations for the individualized education plan (IEP) team to have short-term objectives for each child in his or her IEP will no longer be required as part of the annual goals.
- A child’s progress report toward meeting the annual goals must be reported to the IEP team as in the previous IDEA legislation. With the new law, however, there is no longer a reference to “the extent to which the progress is sufficient to attain the goal by the end of the year.” The amendments clarify that the transition process for a student with a disability now begins at age 16. In the past, only transition planning, but not the actual transition process, would begin at age 16.
- A new section allows IEP team members to be excused from attendance if their area is not being discussed. When this section is applied with new provisions allowing alternate means of meeting participation (eg, conference calls), consolidation of reevaluation meetings and other IEP meetings, and a pilot program authorizing up to 15 states to use multiyear IEPs, the combined effect is a transformation of the traditional IEP meeting that had been a face-to-face meeting that required all participants to sit around a table at the school.
- The Secretary of Education is authorized to approve proposals from up to 15 states to allow local school districts to offer, with parental consent, a multiyear IEP not to exceed 3 years.
- The Secretary of Education is authorized to grant waivers of statutory and regulatory requirements, for a period not to exceed 4 years, to 15 states that propose to reduce excessive paperwork and noninstructional time burdens. The Secretary is prohibited from waiving requirements related to civil rights or the right of a child to a free appropriate public education.
- Parents of a child who is transitioning from part C (early childhood) to part B (school-age) services can request that an invitation to the initial IEP meeting be sent to representatives of the part C system to assist with a smooth transition of services. This provision does not require a part C representative to attend, but it does encourage collaboration.
- Services comparable to those described in the IEP that are in effect before a child’s transfer to a new school must be provided by the new school district. These services must continue until the previous IEP is adopted or a new IEP is developed, adopted, and implemented; regardless of whether the child is transferring in the same state or from another state, the child’s previous IEP will be valid. This new provision will help parents of transferring students know what they can expect from their new schools.
- The procedural safeguards notice, which explains the specific rights and responsibilities of the parent in the special education process, will be routinely distributed only once per year. However, a copy will be distributed after the initial referral, when a parent makes a request for an evaluation, when a due process complaint has been filed, or if a parent requests a copy.
Parents now have 2 years in which to exercise their due process rights after they knew or should have known that an IDEA violation has occurred. Other due process changes can be found at www.pacer.org/idea/2004/summary.htm. The due process hearing is an impartial procedure used to resolve disagreements over issues related to special education services that arise between a parent and a school division. The right of the parents or the school division to request a due process hearing is guaranteed by federal and state laws that govern the education of children with disabilities.

The right of a student with a disability to stay in his or her current educational placement pending an appeal is eliminated for alleged violations of the school code that may result in a removal from the student’s current educational placement for more than 10 days. Before this update of the IDEA, the student with a disability would have been allowed to stay in his or her current educational placement pending an appeal regardless of how many days the violation would have removed him or her from the current placement.

A child is entitled to receive programming and services necessary to enable him or her to receive a free appropriate public education consistent with section 612(a)(1) during the period in which he or she is in an interim alternative education setting.

Before the IDEA 2004, the burden was on the school district to show that the behavior resulting in a disciplinary action was not a manifestation of the child's disability before being allowed to apply the same disciplinary procedures as it used for nondisabled children. Other changes in discipline can also be found at www.pacer.org/idea/2004/summary.htm.

Medically Necessary Versus Educationally Needed
Health care professionals frequently view educationally related services as medically necessary or helpful for children and adolescents with chronic diseases and disabling conditions. Although this is appropriate in the health care setting, it is not the standard for services mandated to be provided by public education systems. The additional proviso that the service must be necessary for education or special education is a key component in the laws. Related services are those services indicated as necessary for the child to maximize his or her special education program (ie, IEP). In other words, without the related services, the child might not be able to maximize his or her special education program. This difference in perspective and interpretation by physicians and other health care professionals and parents often leads to misunderstandings, frustrations, conflicts, and problems in the development and implementation of related services within school programs for children with disabilities. To best serve children with disabilities and their families, physicians and other health care professionals need to be familiar with these issues, their legal basis, and the special educational process and system. Maintaining this knowledge is a key function of the medical home provider for children with chronic diseases and disability conditions. Readers are referred to the American Academy of Pediatrics policy statement "The Pediatrician’s Role in Development and Implementation of an Individual Education Plan (IEP) and/or an Individual Family Service Plan (IFSP)

In addition, functional classifications as suggested by the World Health Organization in its International Classification of Impairments, Disabilities, and Handicaps offer many advantages to the current diagnostic systems used by the medical home provider for children with disabilities. In 2001, the World Health Organization approved the International Classification of Functioning, Disability, and Health as the international standard for conceptualizing, classifying, and coding function. It evaluates all children within the same structure and metric regardless of diagnosis. It highlights a child’s unique strengths and needs for the purpose of directing interventions. This is particularly advantageous in the case of spectrum diagnoses such as autism and cerebral palsy in which a label alone is not sufficient to direct service interventions. A functional assessment of the child provides a more complete picture so that providers can better match resources and needs. Functional classification also looks at individuals within the context of their social and physical environment, taking into account the impact of human and technologic supports on an individual’s “activities and participation.” In addition, functional classification catalyzes interdisciplinary communication and collaboration by providing a common structure and language for discussion.

Challenges for Schools
Providing related services presents significant opportunities for the children served as well as challenges for the educational system. With greater numbers of children with chronic diseases and disabling conditions entering the school system and the increasing complexity of these conditions, many issues and problems have developed. The availability of services, designation of responsibility for their payment and provision, and conflicting legal imperatives, as well as other obstacles, result in vastly different services in various communities. The current trend of integration and inclusion of many children with a wide range of disabilities in “regular” classrooms and programs is making the provision of related services outside of traditional “special” educational settings a larger and more complex issue. Adequate classroom and school modifications (eg, ramps and accessible sinks and toilets) and support personnel (eg, instructional assistants, school nurses, and special education teachers) are needed in more classrooms and schools.
The difficulties in implementation of related services in schools are as varied and complicated as the disabilities of the children involved. These problems, among others, include:

- lack of clarity about which circumstances should result in a child’s exclusion from school for medical reasons;
- uncertainty about the responsibility for, and administration of, complex nursing treatment or therapy in school;
- inconsistencies in state and local guidelines and interpretations about which health care professionals should prescribe the type and amount of physical, occupational, and speech therapies;
- uncertainty about medical liability for therapies administered in school;
- conflicting opinions about the appropriateness of some therapies used for children;
- concern about the rising cost of special education services and whether all treatment required in IEPs is warranted; and
- the frequent lack of provision of related services for children who may not qualify for special education but who have chronic diseases and disabling conditions that impair their ability and readiness to attend or participate in school.9

School-System Responsibility

In March 1999, the US Supreme Court ruled that complex nursing service (ventilator care) is a related service.10 The difference between educationally related services and rehabilitation services is unclear. Court rulings have generally mandated that all therapies and equipment (eg, assistive devices) recommended in the IEP be reimbursed by the educational system.11 However, this has not precluded the application of Medicaid or other public funding for payment of equipment or to support medical service provisions within the IEP for children with disabilities. Although private insurance carriers have generally declined reimbursement for therapies provided in the schools, in specific situations they may be responsible for payment of school-based services and frequently pay for community-based services. One example of private insurance carriers paying for these services would be during the summer when school is not in session. Even if insurance payment is an option, the parents may decide not to make claims against their insurance, because it would create a threat of financial loss, such as lowering the child’s available lifetime medical benefits. Generally, school systems are not responsible for acute rehabilitation services.

In communities in which the school systems have borne the responsibility for implementing the IEP and funding most of the therapies, the educational authorities are increasingly concerned about the responsibility for overseeing the provision of complex nursing care and other related services for children with disabilities who attend public school. School systems also are concerned about insurance companies and managed care systems shifting funding responsibilities for rehabilitation and medical diagnostic services from health care to the school system. Each state’s mandate to the local school system may vary in the degree that any such services are paid by the school system. The variability of school systems to assume responsibilities has the potential to (1) increase conflicts with local physicians and other agencies responsible for health care provision, (2) contribute to the disjointed nature of health care for children, and (3) result in unnecessary treatment at increased cost,12 which depletes educational resources for other children.

The special needs of students with complex health conditions that require modifications in the school environment are also commonly documented in an individual student health plan, also known as a student services plan, nursing care plan, or student medications plan. Although these plans are not mandated by law in every state, such plans typically provide information on a student’s chronic health condition, instructions on the administration of medication, and emergency contact information. A combination of IEP and individual student health plan is often necessary to help manage a student’s health condition in the school and classroom settings.

The Physician’s Role

The physician’s role mandated by the IDEA as a related service is defined to include only the diagnosis and evaluation of the disability. However, in the context of the medical home, the physician’s role also includes the medical management, supervision, and program planning for the child. The IDEA does not mandate that these additional roles be paid for by the public school. Parents often need an advocate for the child who can be objective in assessing the child’s special needs and determining realistic expectations. Input from the medical home professionals also assists with placing services in a developmental context in which changes in needs are to be expected over time. The important medical services extend beyond IDEA mandates. Currently, the funding for the physician and other health care professionals’ time to complete this role is lacking in most health insurance programs and is not funded by most federal and state funds for education. However, physicians can bill for their consultative services and for other related services with some private insurance plans, Medicaid, and the State Children’s Health Insurance Program.

CONCLUSIONS

A multidisciplinary assessment within the school system is required in the initial evaluation of children to deter-
mine their eligibility for services within the educational system. It is also necessary to maintain a comprehensive multidisciplinary approach in the provision of these services, which should be coordinated with the child’s medical home professional. The inequalities in the interpretation and provision of services between and within states and school districts need to be corrected. The developmental, educational, functional, and medical needs of the child or adolescent should be determined first, and then the appropriate services to meet these needs should be provided in a timely manner. Issues of who provides the appropriate services and how payment is to be made must be resolved in the context of maintaining the child in the appropriate educational environment.

RECOMMENDATIONS

1. Physicians and other health care professionals should be well informed about the medical and educational needs of children and adolescents with chronic diseases and disabling conditions.

2. Educational opportunities should be developed and made available to physicians and other health care professionals at local, state, and national levels.


4. Pediatricians, including pediatric subspecialists, and other health care professionals should objectively appraise the special needs of children and adolescents, determine realistic expectations, and advocate for children and adolescents by assisting with establishing an appropriate balance between the recommendations made by the school team and the desires of the family.

5. The initial pediatric focus for services should be on the child or adolescent with a disability and on his or her specific needs, and these needs should not necessarily determine the child’s placement. Once these specific needs have been defined, the role of the school system and the role of community providers should be determined. The specific class placement should not determine the provision of related services in school.

6. Care coordination for children and adolescents with chronic and disabling conditions should take place in the medical home, and the medical home must include the primary care physician, pediatric specialists, and other health and human services professionals regardless of the location of, or source of payment for, these services.

7. Physicians and other health care professionals should take a more active role in the development and implementation of individualized family service plans.

8. Physicians and other health care professionals should get involved at the systems level. Physicians, especially pediatricians, should seek representation on local advisory and interagency committees that oversee programs for school placement of children and adolescents with chronic diseases and disabling conditions.

COUNCIL ON CHILDREN WITH DISABILITIES EXECUTIVE COMMITTEE, 2006–2007
Paul H. Lipkin, MD, Chairperson
*J. Daniel Cartwright, MD
Larry W. Desch, MD
John C. Duby, MD
Ellen Roy Elias, MD
Chris Plauché Johnson, MD, MEd
Eric B. Levey, MD
Gregory S. Liptak, MD
Nancy A. Murphy, MD
Scott M. Myers, MD
Ann Henderson Tilton, MD

LIAISONS
Donna Gore Olsen
Family Voices
Donald Lollar, EdD
Centers for Disease Control and Prevention
Michelle Macías, MD
Section on Developmental and Behavioral Pediatrics
Bonnie Strickland, PhD
Maternal and Child Health Bureau

STAFF
Stephanie Mucha Skipper, MPH

*Lead author

REFERENCES

4. Individuals With Disabilities Education Improvement Act of 2004, HR 1350, November 19, 2004
6. American Academy of Pediatrics, Committee on Children With Disabilities. The pediatrician’s role in development and implementation of an individual education plan (IEP) and/or an individual family service plan (IFSP). Pediatrics. 1999;104:124–127
7. World Health Organization. International Classification of Impair...

SUGGESTED READING
Ordover EK, Boundy KB. Educational Rights of Children With Disabilities: A Primer for Advocates. Cambridge, MA: Center for Law and Education; 1991

INTERNET RESOURCES
Utah MedHome Portal. Education issues. Available at: http://medhome.med.utah.edu/education/index.cfm
KidSource Online. Overview of ADA, IDEA, and Section 504. Available at: www.kidsource.com/kidsource/content3/ada.idea.html
National Dissemination Center for Children With Disabilities. The education of children and youth with special needs: what do the laws say? Available at: www.nichcy.org/pubs/newsdig/nd15txt.htm
National Early Childhood TA Center. IDEA reauthorization 2004. Available at: www.nectac.org/idea/idea.asp#summaries
Council for Exceptional Children. The new IDEA. Available at: www.cec.sped.org/AM/Template.cfm?Section=Search&section=Policy_and_Advocacy1&template=/CM/ContentDisplay.cfm&ContentFileID=723
American Academy of Pediatrics. School health resources. Available at: www.schoolhealth.org
Discover IDEA Web site. IDEA ’97 final regulations: related services. Available at: http://specialed principals.org/discidea/regs/definitions.htm#sec300.24
Maddux CD. Some special education sites you may wish to visit. Available at: http://unr.edu/homepage/maddux/splinks.html
UCLID Center at the University of Pittsburgh. Functional classification. Available at: www.uclid.org.8080/uclid,functional.classification.html

ADVOCACY SUPPORT RESOURCES
Family Voices. Available at: www.familyvoices.org
Technical Assistance Alliance for Parent Centers. Available at: www.taalliance.org
Parent Advocacy Coalition for Educational Rights. Available at: www.pacer.org
Families and Advocates Partnership for Education. Facts on hand: related services. Available at: www.fape.org/pubs/fape.33.pdf
Council for Exceptional Children. Available at: www.cec.sped.org
Learning Disabilities Association of America. Available at: www.lddaamerica.org
National Center for Learning Disabilities. Available at: www.ncld.org
Office of Special Education and Rehabilitative Services. Available at: www.ed.gov/about/offices/list/osers/index.html?src=oc
STATE CHILDREN’S HEALTH INSURANCE PROGRAM ACHIEVEMENTS, CHALLENGES, AND POLICY RECOMMENDATIONS

Committee on Child Health Financing

ABSTRACT
This policy statement reviews the impressive progress of the State Children’s Health Insurance Program since its enactment in 1997 and identifies outstanding challenges and state and federal policy recommendations. The American Academy of Pediatrics urges Congress to reauthorize SCHIP to strengthen its historic gains. The following set of recommended strategies for reauthorization pertain to funding, eligibility and enrollment, coverage, cost sharing, payment and provider-network capacity, and quality performance.

INTRODUCTION
The State Children’s Health Insurance Program (SCHIP), enacted in 1997 as Title XXI of the Social Security Act (Pub L No. 105–33), has achieved remarkable progress in its brief history. As a result of SCHIP, health insurance has been extended to millions of children from low-income families, and rates of uninsurance among this population have declined by 2.2 million, from 23% in 1997 to 14.4% in 2004.1 Access to health care has been vastly improved.2 Specifically, because of SCHIP, more children have a medical home, more children receive preventive care and immunizations, and fewer children have an unmet need for dental care.2–7 Family satisfaction and quality of care have also improved significantly under SCHIP.4,7 Income and racial/ethnic gaps in health insurance coverage and access to care have also narrowed.8 SCHIP also has had positive spillover effects on the Medicaid program.9 As a result of SCHIP outreach, millions of potentially eligible but uninsured children have been enrolled in Medicaid.10 Eligibility-determination processes have been simplified, and coordination between SCHIP and Medicaid has become increasingly effective.3 The landmark SCHIP legislation allowed states to design their SCHIP programs as expansions of Medicaid, as separate non-Medicaid programs, or as combinations of the two. Unlike Medicaid, SCHIP is not an entitlement. It is capped at the amount that is funded by Congress and the states. States were able to pursue different approaches for offering the most comprehensive, affordable coverage possible for near-poor children and their families. SCHIP is important now more than ever because of concerns about the increased numbers of children with obesity, diabetes, mental health disorders, asthma, and other chronic conditions and the importance of ensuring that these children will be given timely and continuous access to health care services over the span of childhood and adolescence.

Despite the program’s widely acknowledged success and popularity, several
outstanding challenges have been identified by SCHIP officials, enrolled families, participating pediatricians and other health care professionals, and health service researchers. These challenges pertain to (1) ensuring adequate funding, (2) extending the reach of SCHIP to all potentially eligible children and to more uninsured children and families at higher income levels, (3) improving benefit coverage in non-Medicaid plans, (4) maintaining affordable premiums and other forms of cost sharing, (5) providing adequate payments and strengthening provider-network capacity, and (6) improving quality performance. This policy statement identifies recommended strategies in each of these 6 areas, which the American Academy of Pediatrics (AAP) believes will further the program’s success in the next decade.

BACKGROUND
In 2005, SCHIP programs provided health insurance to 4 million children nationwide.11 States selected different approaches to provide health insurance under SCHIP; 21 states created a combination Medicaid and non-Medicaid program, 18 states created a non-Medicaid program, and 17 states and territories and the District of Columbia created a Medicaid program.12 In 27 states and the District of Columbia, eligibility levels are established at the congressional target of 200% of the federal poverty level (FPL), and in 13 states, eligibility has been extended to children with family incomes above 200% of the FPL. Eligibility extends up to 300% of the FPL in 5 states (Connecticut, Maryland, Missouri, New Hampshire, and Vermont) and 350% of the FPL in 1 state (New Jersey).

The original funding-allocation formula for SCHIP, which will expire in 2007, is based on each state’s share of low-income children, its share of low-income uninsured children, and the state’s cost of providing health care services. Funds not spent by states within an allotted time are redistributed to other states according to a specific formula. Unfortunately, in fiscal year 2007, 17 states face SCHIP funding shortfalls that amount to approximately $1 billion according to the Center on Budget and Policy Priorities.13 Shortfalls occurred because of the size of the population of uninsured children, the growth in the population of children from low-income families, the growing instability of employment-based health insurance, and inflation.

In addition to the very serious federal budget shortfalls, since 2001 states have experienced significant budget shortfalls that have adversely affected their ability to sustain their SCHIP programs. The most common cost-cutting response has been to limit outreach and enrollment; few states have actually lowered eligibility or benefits or imposed significantly higher cost-sharing requirements.14 These cost-cutting actions resulted in a first-ever dip in enrollment in 2003.15

The scope of coverage for SCHIP programs in the 39 states that are offering a non-Medicaid plan to some or all of their SCHIP enrollees, although not as comprehensive as Medicaid coverage, still (with few exceptions) far exceeds benefits in employer-sponsored health insurance plans.16 Similarly, although premium rates, copayments, and other dollar limits impose financial burdens for some families, they are still markedly less than those in private health insurance plans, and families consider them reasonable and affordable.17

Provider payment rates, however, are generally low—well below commercial rates—and in many states are at the same level as Medicaid rates. Medicaid professional fees were estimated to be approximately 70% of Medicare rates in 2004 according to the 2006 AAP Pediatric Medical Cost Model developed by actuaries at Reden & Anders.18

The AAP recommends the following improvements to strengthen SCHIP:

1. Ensure adequate funding
   - Establish a new funding-formula approach that relies on a combination of national and state data that does not penalize states for successfully enrolling uninsured children, that takes into account state variations in the costs of providing care, and that extends the period during which redistributed funds can be spent.
   - Set the budget baseline for SCHIP at a rate significantly higher than the level set in law for the final year of SCHIP’s initial authorization to avoid future budget shortfalls.

2. Extend eligibility and enrollment
   - Establish a performance-based outreach fund that rewards states that are more successful in enrolling uninsured children who are eligible for public coverage.
   - Continue to improve on administrative simplification to facilitate enrollment and reenrollment, including shortened forms, streamlined verification requirements, online enrollment, and renewal assistance. In addition, grant states the flexibility to automatically enroll children into SCHIP (and Medicaid) on the basis of findings of other meanstested programs such as the Supplemental Nutrition Program for Women, Infants, and Children (WIC), the National School Lunch Program, and the Food Stamp Program.
   - Encourage presumptive eligibility for all children by allowing health care professionals and designated agencies to grant eligibility for up to 60 days while a child goes through the enrollment process. In addition, encourage states to adopt 12-month continuous eligibility for SCHIP-enrolled (and Medicaid-enrolled) children.
• Allow households with children in both Medicaid and SCHIP to enroll in the same program to ensure continuity among siblings with their pediatric medical home.

• Encourage expansion of SCHIP to include adolescents 19 through 21 years of age and allow emancipated minors eligibility for SCHIP on the basis of their own income. In addition, eliminate eligibility restrictions for dependents of state employees if they qualify on the basis of income.

• Encourage higher income eligibility levels (>200% of the FPL) and discontinue the practice of counting family assets to extend eligibility to more uninsured children.19

• Offer SCHIP buy-in options for children whose family incomes are above their state’s SCHIP eligibility level but who do not have access to or cannot afford comprehensive private health insurance.

• Allow states to cover legal immigrant children who enter the United States on or after August 1996. These children, under the 1996 Welfare Law, are ineligible for Medicaid and SCHIP coverage during their first 5 years in the United States. Other complex rules restrict legal immigrant children from gaining public coverage until they are citizens.

• Allow states to draw down Medicaid/SCHIP matching funds when employers pay for a share of the cost of coverage for children of low-income families enrolled in Medicaid or SCHIP.

• Encourage waiver applications of the Centers for Medicare and Medicaid Services to expand SCHIP coverage for uninsured pregnant women and parents if states have already maximized comprehensive coverage and full enrollment of children.

3. Support comprehensive coverage

• Preserve Medicaid benefit coverage in states with Medicaid SCHIP programs.

• Encourage states to adopt SCHIP benefit packages that are consistent with the AAP policy statement “Scope of Health Care Benefits for Children From Birth Through Age 21,”20 including oral health services, the full range of mental health services, and substance abuse treatment. Preventive care, immunization standards, and periodicity schedules also should be consistent with current AAP requirements. In addition, definitions of medical necessity should adhere to AAP recommendations.21

• Extend eligibility for the Vaccines for Children Program to all children enrolled in non-Medicaid SCHIP programs.

• Eliminate the prohibition against partial benefit packages to allow states with non-Medicaid SCHIP programs to provide additional wrap-around coverage to children, especially those with special health care needs who have inadequate private health insurance.

4. Maintain affordable coverage

• Eliminate differences in copayments and coinsurance for physical and mental health services.

• Adopt cost-sharing policies that do not shift cost to pediatricians, hospitals, and other health care professionals and do not deter the use of medically necessary services. Deductibles and coinsurance should not be used; rather, cost sharing should be in the form of income-adjusted premiums and copayments.

• Maintain policy that requires all preventive services under SCHIP to be exempt from cost sharing.

5. Improve provider payments and network capacity

• Establish payment rates under SCHIP for pediatric services that are at least equal to the most current Medicare RBRVS (Resource-Based Relative Value Scale) rates.

• Ensure adequate payment when new vaccines and other new technologies are introduced. Under capitated arrangement, states should ensure that provisions are made to reimburse physicians for all vaccine-related overhead costs (vaccine product-acquisition and administration costs) of the new vaccines until new contracts are negotiated. In addition, physicians should receive payment for the expenses associated with the administration of each vaccine.

• Adopt financial incentives for medical homes, especially in the care of children with special needs, including chronic care management, child and family education, and coordination and consultation with pediatric specialists and other support services.

• Provide financial incentives for pediatric practices that adopt quality-performance goals.

• Recognizing the dearth of pediatric subspecialists nationwide, encourage the inclusion of pediatric subspecialists, and the academic medical centers in which they practice, in managed care plan networks, and encourage coordination and communication between pediatric subspecialists and primary care practitioners.

• Identify new mechanisms to designate and support safety net providers, including office-based pediatric practices and hospitals that specialize in the care of children, who serve a certain proportion of publicly insured children.
6. Strengthen quality performance

- Adopt a consistent conceptual framework (eg, the framework of the Institute of Medicine) to assess health care quality across SCHIP programs that takes into account the unique features of child health and health care. Performance goals for states and the plans with which they contract should consist of short-term and long-term health care outcomes, including monitoring eligibility thresholds and projected enrollment volume, program retention, access to medical care, assessments of process and outcomes of pediatric care, and family and provider satisfaction.
- Improve the collection and analysis of individual-level enrollment data and claims-based utilization data.
- Involve pediatricians, pediatric subspecialists, pediatric mental health professionals, pediatric dentists, and other pediatric clinicians and families, including those who represent special populations, in continuously reviewing and evaluating each state’s SCHIP.
- Expand funding support for SCHIP evaluations and allow greater access to state data for research.
- Measures should be appropriate for children’s health. Any effort to measure quality should take into account the unique features of child health and health care. In addition, pediatric and family representatives should be included in all measurement efforts at the national, state, and local levels.

CONCLUSIONS

SCHIP has a proud history on which to build. To achieve continued success in reducing uninsurance among children and ensuring access to high-quality pediatric care, the AAP recommends that Congress and state policy makers adopt these important recommendations.

COMMITTEE ON CHILD HEALTH FINANCING, 2006–2007
*Steve Wegner, MD, FAAP, Chairperson
Charles Barone, MD, FAAP
Thomas Chiu, MD, FAAP
Anthony Johnson, MD, FAAP
Richard Lander, MD, FAAP
Russell Libby, MD, FAAP
Thomas Long, MD, FAAP
Mark Reuben, MD, FAAP
Corinne Walentik, MD, FAAP

CONSULTANT
*Margaret McManus

STAFF
Teri Salus, MPA, CPC

*Lead authors

REFERENCES

14. Fox HB, Limb SJ. SCHIP Programs More Likely to Increase Children’s Cost Sharing Than Reduce Their Eligibility or Benefits to Control Costs. Washington, DC: Maternal and Child Health Pol-
RESIDENTS’ ATTITUDES ABOUT DUTY HOURS REGULATIONS

“To better understand the perspectives of residents on the effects of the ACGME duty hours restrictions, Jennifer S. Myers, MD, and colleagues performed a multi-site survey of internal medicine and surgery residents, focusing on residents who were in training both before and after implementation of the new regulations. The survey questions were designed to elicit opinions in three areas: quality of patient care and safety, residency education, and quality of resident life. . . . Medical and surgical residents’ opinions of quality of care and medical errors were similar to each other. Both groups of residents felt that the quality of care had decreased slightly after implementation of the new regulations, but that the continuity of care had decreased a great deal. They also felt that errors attributable to continuity of care had increased, but that errors related to resident fatigue had decreased. Residents felt that the new rules had created a ‘shift-work’ mentality among housestaff, but did not believe that the quality of program graduates had changed. In addition, they felt that their quality of life had improved substantially since the implementation of the regulations. . . . The authors note that the survey results indicate that medical errors related to fatigue might have been replaced with errors related to discontinuity of care as a result of duty hours reform. Furthermore, duty hours reform has not resulted in significantly more hours of sleep per week for residents. Residents have also reported reductions in bedside teaching and in opportunities for mentoring from attending physicians. The authors state that these unintended consequences of duty hours reductions will need to be addressed as residency programs adapt their education programs to meet regulatory requirements.”

Myers JS. Academic Physician & Scientist. February 2007
Noted by JFL, MD
Electronic Prescribing Systems in Pediatrics: The Rationale and Functionality Requirements

Council on Clinical Information Technology

ABSTRACT
The use of electronic prescribing applications in pediatric practice, as recommended by the federal government and other national health care improvement organizations, should be encouraged. Legislation and policies that foster adoption of electronic prescribing systems by pediatricians should recognize both specific pediatric requirements and general economic incentives required to speed the adoption of these systems. Continued research into improving the effectiveness of these systems, recognizing the unique challenges of providing care to the pediatric population, should be promoted.

BACKGROUND
The American Academy of Pediatrics (AAP) is committed to providing the best and safest health care system possible for children.

Statement of Problem
The AAP recognizes that the “increasing complexity in patient care in addition to the public’s increased scrutiny of the health care system underscores the need to make patient safety an issue of high priority.”1 The AAP supports national efforts to improve patient safety and the recommendations of the Institute of Medicine, the Institute for Safe Medical Practices, the Leapfrog Group, and others who encourage the implementation and use of electronic prescribing (e-prescribing) by physicians as a method to improve patient safety.2,3

New Information
E-prescribing systems reduce transcription errors by eliminating illegible prescriptions. Computerized decision support can ensure that prescriptions are checked for drug-drug and drug-allergy interactions before the prescription is written. Dosage calculators can ensure that the correct dose of medication is given on the basis of patient age and weight, and dose-range checking can alert prescribers when doses outside the predetermined ranges are prescribed. Many e-prescribing systems can also check formulary information to determine if a selected medication is covered by a patient’s insurance, thereby decreasing patient drug cost and increasing both patient and physician compliance with insurers’ preferred-drug prescription programs.5–9 Additional information is available in the accompanying technical report on e-prescribing.4 Research examining the impact of e-prescribing on reducing malpractice claims might result in a commensurate reduction in malpractice.
liability insurance. Ongoing research will be needed to study the types of errors that may continue to occur after implementation of these systems, including potential new types of errors introduced by the use of e-prescribing systems. This research will guide refinements and improvements to the effectiveness of these systems.6,10,11

SUMMARY/CONCLUSIONS
The AAP believes there is sufficient evidence supporting the ability of e-prescribing systems to prevent medical errors and enhance patient care.5–9 However, as with any new technology, the use of these systems may have unintended consequences or novel risks that will need to be monitored and studied over time.

RECOMMENDATIONS
1. Federally sponsored research should be conducted to document, in both inpatient and ambulatory (office) settings, specific characteristics of e-prescribing systems that are most beneficial in preventing errors and enhancing patient care. Office processes and methods of implementation that facilitate the effective and efficient use of e-prescribing systems require study.10,12–14

Accurate data on the incidence and scope of prescribing errors, adverse drug events, and near-miss errors must be available. Regulations should be promoted to facilitate no-fault, anonymous adverse drug event reporting systems as an enabling step toward understanding and intervening to prevent medical errors.1

2. Because safety for children is paramount, e-prescribing systems used for the care of children should include, at a minimum, pediatric-specific decision support such as weight-based dose calculations and alerts and pediatric drug information and formulation options.3,7,15–17

When possible, e-prescribing systems should be implemented as part of a robust electronic health record. Such implementations offer advantages well beyond those of freestanding e-prescribing systems. When implementing a stand-alone e-prescribing system, thought should be given to the potential future need to transfer data to, or interface the e-prescribing system with, an electronic health record.

3. Federal legislation that would unify state regulations and allow for e-prescribing and digital transmission of all prescriptions directly to pharmacies, including those for controlled drugs, should be encouraged.18 The AAP furthermore supports legislation that would require all pharmacies, either directly or through a clearinghouse, to accept digitally transmitted and signed prescriptions. The AAP supports a process for the development of standards for the transmission of digital prescriptions, analogous to the standards-development process under the Health Insurance Portability and Accountability Act for electronic data interchange.

4. Despite significant benefits to medical and liability insurers, patients, and pharmacy benefit managers,19 e-prescribing applications are an office-practice expense that generates a disproportionately small or no pediatric practice revenue; therefore, the AAP believes adoption of e-prescribing technology would be hastened by the offering of incentives such as pay-for-performance bonuses to practices that routinely use e-prescribing systems that incorporate clinical decision-support alerts.

5. Because practitioners in rural or low-income areas may face financial and system barriers and, in many cases, do not have access to the network infrastructure to support e-prescribing systems, federal grant and loan programs should be available to support system enhancements such as Internet access and start-up costs.

IMPLEMENTATION
Recommendations 1 and 5 (federally funded research and federal grants and loans for e-prescribing systems) may be implemented by providing research grants through the National Library of Medicine, the Agency for Healthcare Research and Quality, the Health Resources and Services Administration, and other federal and local agencies.

Recommendation 2 (minimum standards for e-prescribing systems) may be implemented by educating providers before purchase of such systems on the required elements through published reports such as the accompanying technical report.4 Such reports should also be shared with standards-development organizations to encourage the inclusion of minimum requirements into the development of these standards.

Recommendation 3 (federal legislation on e-prescribing) requires action by the collaborative action of the Drug Enforcement Administration to develop standards for the secure digital transmission of category II controlled substances and enable federal legislation that takes precedence over the restrictions placed by state regulations.

Recommendation 4 (incentives for purchase) should be part of federal and state initiatives to reduce medical errors. Efforts to encourage larger insurers to underwrite such systems should continue—with demonstration projects to document the cost savings to them by the adoption of e-prescribing systems.

COUNCIL ON CLINICAL INFORMATION TECHNOLOGY, 2006–2007
Mark M. Simonian, MD, Chairperson
Joseph H. Schneider, MD, MBA, Vice Chairperson
Kristin A. Benson, MD
Donna M. D’Alessandro, MD
Mark A. Del Beccaro, MD
Willa H. Drummond, MD, MS
George R. Kim, MD
Michael Leu, MD
Gregg C. Lund, DO
Eugenia Marcus, MD
Alan E. Zuckerman, MD

PAST COUNCIL EXECUTIVE COMMITTEE MEMBERS
S. Andrew Spooner, MD, MS, Immediate Past Chairperson
*Robert S. Gerstle, MD
Kevin B. Johnson, MD, MS
Christoph U. Lehmann, MD

STAFF
Rebecca Marshall

*Lead author

REFERENCES
CLINICAL REPORT

Evaluation of Suspected Child Physical Abuse

Nancy D. Kellogg, MD, and the Committee on Child Abuse and Neglect

ABSTRACT
This report provides guidance in the clinical approach to the evaluation of suspected physical abuse in children. The medical assessment is outlined with respect to obtaining a history, physical examination, and appropriate ancillary testing. The role of the physician may encompass reporting suspected abuse; assessing the consistency of the explanation, the child’s developmental capabilities, and the characteristics of the injury or injuries; and coordination with other professionals to provide immediate and long-term treatment and follow-up for victims. Accurate and timely diagnosis of children who are suspected victims of abuse can ensure appropriate evaluation, investigation, and outcomes for these children and their families.

PREVALENCE
In 2004, 152,250 children and adolescents were confirmed victims of physical abuse in the United States.1 Of the 4 types of child maltreatment (neglect, physical abuse, sexual abuse, and emotional abuse), physical abuse is second to neglect, constituting approximately 18% of the total.1

Despite these statistics, the estimated number of victims is much higher; in 1 retrospective cohort study of 8,613 adults, 26.4% reported they were pushed, grabbed, or slapped; had something thrown at them; or were hit so hard they got marks or bruises at some time during their childhood.2 It has been estimated that 1.3% to 15% of childhood injuries that result in emergency department visits are caused by abuse.3 Physical abuse remains an underreported (and often undetected) problem for several reasons including individual and community variations in what is considered “abuse,” inadequate knowledge and training among professionals in the recognition of abusive injuries, unwillingness to report suspected abuse, and professional bias. For example, in 1 study,4 31% of children and infants with abusive head trauma were initially misdiagnosed. Misdiagnosed victims were more likely to be younger, white, have less severe symptoms, and live with both parents when compared with abused children who were not initially misdiagnosed. Such studies suggest a need for practitioners to be vigilant to the possibility of abuse when evaluating children who have atypical accidental injuries or obscure symptoms that are suggestive of traumatic etiologies but who do not have a history of trauma.

Child abuse has significant long-term medical and mental health morbidity.5 Children with abusive head or abdominal injuries are more likely to die or become more severely incapacitated than are children with head or abdominal injuries caused by accidents.6-8 Victims of physical abuse in childhood are more likely to
develop a variety of behavioral and functional problems including conduct disorders, physically aggressive behaviors, poor academic performance, and decreased cognitive functioning. Additional problems include anxiety and depression, as well as social and relationship deficits.

CHARACTERISTICS OF VICTIMS AND RISK FACTORS
Child physical abuse affects children of all ages, genders, ethnicities, and socioeconomic groups. Male and female children experience similar rates of physical abuse. In 1 survey study of more than 2000 children and adolescents, 15% of adolescents received injuries from a physical assault and were more likely than children in younger age groups to receive injuries from abuse. Although the risk of physical abuse increases with age, fatal abuse and serious abusive injuries are more common among children and infants younger than 2 years. Children in homes with annual incomes of less than $15,000 per year have 3 times the number of fatalities, 7 times the number of serious inflicted injuries, and 5 times the number of moderate inflicted injuries when compared with children living in homes with annual incomes of greater than $15,000 per year. Risk factors for infant maltreatment include maternal smoking, the presence of more than 2 siblings, low infant birth weight, and an unmarried mother. One study found that children living in households with unrelated adults were approximately 50 times more likely to die of inflicted injuries than were children residing with 2 biological parents. The US Department of Health and Human Services has indicated that the rate of physical abuse is 2.1 times higher among children with disabilities than children without disabilities. The presence of risk factors should not be used as indicators of child abuse but rather to provide guidance in prevention strategies as well as management and treatment plans.

ROLE OF THE PEDIATRICIAN
The role of the pediatrician encompasses prevention of abuse and detection and medical management of victims of abuse. Accurate identification of children who are suspected victims of abuse can facilitate appropriate evaluation, referral, investigation, and outcomes for these children and their families. Children usually sustain abuse at the hands of a caregiver who misinterprets and responds inappropriately to the child’s behavior. For example, caregivers who had smothered, shook, or slapped their infant within the first 6 months of life were more likely to be worried about crying and to believe that their infants cried excessively. There is a close correlation between the age-specific incidence curve of infants hospitalized with abusive head trauma and the age-specific normal crying behavior of infants up to 36 weeks of age.

In an anonymous telephone survey of 1435 mothers, 2.6% of children younger than 2 years were shaken by their mothers as a means of discipline. Caregivers may respond inappropriately to their child’s behavior when they are unduly stressed. Poverty, significant life events, and caregiver role conflicts are stressors that are often associated with abuse. Pediatricians can effectively educate parents regarding the range of normal behaviors in infants and children, provide anticipatory guidance, and be a resource when the behavior becomes unmanageable for parents. In addition, pediatricians can screen for adult-partner violence; in 1 study, child abuse was 4.9 times more likely in families with identified spouse abuse than in families without identified spouse abuse. Other conditions that place children at risk of being abused, such as maternal depression or drug abuse, may also be identified.

Careful medical assessment, detection of suspicious injuries, and reporting of abuse may prevent further abusive trauma in infants and adults. In 1 study of abuse victims younger than 24 months, 75% had evidence of previous trauma or history of a previous injury. Child abuse may recur 35% of the time without appropriate detection and intervention.

As with other types of child maltreatment, there have been recent advances in medical knowledge regarding physical abuse. Most recent developments have addressed more accurate differentiation between inflicted and accidental injuries as well as detecting conditions that may mimic abusive injuries. Although consideration of nonabusive causes of injuries may merit additional evaluation and testing, the physician is mandated by law to report suspicions of abuse and should not delay reporting pending confirmatory testing or information. In all states, the law also provides some type of immunity for good-faith reporting. Once a suspected victim is identified and further assessment and management is required, using a pediatric child abuse consultant, if available, early in this process may obviate the need for invasive or expensive testing and can help direct the pediatrician toward appropriate evaluation. The detection and diagnosis of child physical abuse depends on the clinician’s ability to recognize suspicious injuries, conduct a careful and complete physical examination with judicious use of auxiliary tests, and consider whether the caregivers’ explanation is supported by the characteristics of the injury or injuries and the child’s developmental capabilities. The physician should also ensure that the child’s immediate medical and safety needs are met. Child abuse injuries, particularly traumatic brain injuries, may result in significant long-term disabilities including learning deficits, attention-deficit/hyperactivity disorder, behavioral problems, seizures, spasticity, blindness, paralysis, and mental retardation. Continuity of care for such children is essential, especially if they are transferred to other caregivers or foster homes.

Many hospitals and communities have developed
child abuse—assessment teams of pediatricians and other professionals who specialize in the assessment of suspected victims of child abuse.25 Such teams usually have access to additional information from law enforcement and child protective services, such as scene investigation, that may facilitate more thorough injury assessment and diagnosis. Involving such teams early in the process can ensure accurate and comprehensive assessments and information sharing among the medical and nonmedical disciplines involved and can provide for intermediate and long-term management of the child and family. Pediatricians with expertise in evaluating suspected abuse should provide training and assistance to emergency physicians and other first responders to enhance detection and appropriate referral of these patients.

Many regions do not have specialized child abuse teams but do have physicians with expertise in child abuse. Pediatricians should know which hospitals in their region have the most available expertise in the emergency evaluation of suspected child abuse. In turn, pediatricians with expertise in child abuse often act as consultants for emergency departments and child protective services. Close collaboration is necessary, particularly for establishing how the child should be transported between facilities, who should notify child protective services, who should notify the caregiver(s) of suspected abuse and when, and whether law enforcement should be notified. For those who do not require emergent transportation by ambulance, child protective services may facilitate transportation of a suspected child victim from one facility to another, assist in notifying the caregivers and law enforcement of suspected abuse, and provide an emergent safety plan on hospital discharge or clinic dismissal.

DEFINITIONS
The recognition and reporting of physical abuse is hindered by the lack of uniform or clear definitions. Many state statutes use words such as “risk of harm,” “substantial harm,” “substantial risk,” or “reasonable discipline” without further clarification of these terms. Many states still permit the use of corporal punishment with an instrument in schools; on the other hand, the American Academy of Pediatrics has proposed that “striking a child with an object” is a type of physical punishment that “should never be used”26 and has recommended that corporal punishment be abolished in schools.27 The variability and disparities in definitions may hinder consistent reporting practices.

CLINICAL PRESENTATIONS AND SETTINGS
Most physical abuse injuries are likely to not be detected or reported. Minor injuries may not require medical attention and may be obscure or hidden. Infants and children are reported as suspected victims of physical abuse when 1 or more of the following occurs: an individual (including a professional) sees and reports a suspicious injury; an individual witnesses an abusive event; a caregiver observes symptoms and brings the child in for medical care but is unaware that the child has sustained an injury; an individual asks a child if he or she has been hurt in an abusive way; the abuser thinks the inflicted injury is severe enough to require medical attention; or the child victim discloses abuse. The American Academy of Pediatrics has indicated that “hospitalization of children requiring evaluation and treatment for abuse or neglect should be viewed by third-party payors as medically necessary.”28

The clinical approach to an infant or child with possible abusive injuries is not significantly different from standard pediatric care. As with all patients, a severely injured child must be stabilized before further evaluation is undertaken. This initial evaluation may encompass a trauma response team and pediatric specialists in surgery, emergency medicine, and critical care. Careful documentation may not be possible initially and must always be secondary to resuscitation and stabilization of the patient. Once the child is stabilized, a careful and well-documented history, as always, is the most critical element of the medical evaluation. Using quotes whenever possible, the pediatrician should document descriptions of the mechanisms of injury or injuries, onset and progression of symptoms, and the child’s developmental capabilities. The physical examination should include detailed documentation, either by body diagrams and/or photographs, of any concerning cutaneous findings and should include a thorough search for other signs that may suggest a nontraumatic cause. If the child is verbal, it may be helpful to gather parental and patient histories separately. If abuse is a concern after this preliminary evaluation, consultation with a child abuse pediatrician, pediatric specialist, or pediatrician experienced in this area, if available, may be helpful in determining the best way to proceed with assessment.

Physical discipline is commonly inflicted on areas of the body that are concealed by clothing (e.g., back/buttocks). When inflicted injuries are visible or incidentally discovered, child victims and their abusers typically explain the injuries as accidental; if clinicians or professionals are not critical or skeptical of this information, the injuries may be incorrectly attributed to accidental causes. Other victims present with severe inflicted injuries that require medical care. The initial history is typically vague and/or benign and may become inconsistent as the investigation progresses.

MEDICAL HISTORY
The interview of parents or caregivers of infants or children who present with serious injuries may be conducted in an outpatient or inpatient setting. If the child presents to a clinic with a serious injury that requires further medical care in a specialty (e.g., orthopedics) or
hospital setting, the clinician may opt to gather the minimum information to establish a need for reporting to child protective services. Any statements made by the caregiver regarding the injury should be documented accurately and completely. Once the clinician has assessed all the injuries, including approximate ages of injuries (when possible), a careful, complete, and detailed history should be obtained from the caregivers. Explanations that are concerning for intentional trauma include:

1. no explanation or vague explanation for a significant injury;
2. an important detail of the explanation changes dramatically;
3. an explanation that is inconsistent with the pattern, age, or severity of the injury or injuries;
4. an explanation that is inconsistent with the child’s physical and/or developmental capabilities; and
5. different witnesses provide markedly different explanations for the injury or injuries.

Information regarding the child’s behavior before, during, and after the injury occurred, including feeding times and levels of responsiveness, should be gathered. Victims of significant trauma usually have observable changes in behavior. Access to caregivers and caregiver activities before, during, and after the injury occurred are also important to document. Frequently, infants and children present to medical settings with a history of a fall. Recent studies have indicated that short falls may result in bruising; however, more significant types of head trauma, including skull fractures, are exceedingly uncommon but possible.\(^2\)\(^9\)\(^3\)\(^0\)

Information should be gathered in a nonaccusatory but detailed manner. Other information that may be useful in the medical assessment of suspected physical abuse includes:

1. past medical history (trauma, hospitalizations, congenital conditions, chronic illnesses);
2. family history (especially of bleeding, bone disorders, and metabolic or genetic disorders);
3. pregnancy history (wanted/unwanted, planned/unplanned, prenatal care, postnatal complications, postpartum depression, delivery in nonhospital settings);
4. familial patterns of discipline;
5. child temperament (easy to care for or fussy child);
6. history of past abuse to child, siblings, or parents;
7. developmental history of child (language, gross motor, fine motor, psychosocial milestones);
8. substance abuse by any caregivers or people living in the home;
9. social and financial stressors and resources; and
10. violent interactions among other family members.

**PHYSICAL EXAMINATION**

Most injuries of childhood are not the result of abuse or neglect. Minor injuries in children are exceedingly common. Physicians must also consider that unusual events, including accidents, do happen to children\(^3\)\(^1\) and may produce injuries that are not characteristically seen from accidental causes. An injury pattern is rarely pathognomonic for abuse or accident without careful consideration of the explanation provided. In addition, both inflicted and accidental injuries may be seen simultaneously in a child.

**General Assessment**

The child’s alertness and demeanor may reflect neurologic status and degree of discomfort and pain. A thorough and complete neurologic examination must be performed. For example, if alertness appears compromised, eye-opening, verbal, and motor responses should be assessed systematically. Spontaneous and symmetrical movement of all extremities should be noted, as well as any of the child’s responses that indicate pain when extremities are examined and moved. Because abusive caregivers are rarely informative regarding the injuries that have been inflicted, special care should be taken during the examination of the child’s extremities and neck, which may be fractured and require immobilization until diagnostic radiographs can be performed. Evidence of spinal cord injury, such as abnormal reflexes, muscle tone, or responsiveness to tactile stimuli, should be carefully pursued.

When the child is stable, height, weight, and fronto-occipital circumference should be carefully measured and then plotted on a growth chart. Previous measurements obtained from past medical visits should also be obtained to gauge whether growth velocity has been appropriate. Plotting parameters is essential, because clinicians may miss significant growth failure in infants and children if the clinician relies only on their clinical impressions. Physical abuse and failure to thrive are sometimes concurrent\(^1\)\(^2\)\(^-\)\(^3\)\(^3\); in addition, some children are starved intentionally.\(^3\)\(^4\)

Evidence of neglect may be seen during the general examination of the infant or child; extensive dental caries, severe diaper dermatitis, or neglected wound care may be noted in addition to injuries that raise suspicion of abuse. Bald areas on the scalp may sometimes be seen with severe nutritional deficits or with traumatic alopecia. These findings should be differentiated from non-abusive or benign causes such as tinea capitis, alopecia areata, and occipital bald spots caused by supine positioning of young infants.

If the child can be interviewed, his or her demeanor
should be noted during questioning. Some children display strong nonverbal cues of anxiety and reluctance when answering questions regarding potential abuse, because they are protective of their abuser or they fear retribution for “telling.” Others may appear openly fearful of their abuser. Such responses may be important to consider when a safety plan for the child is made.

**Skin Injuries**

Location, size, and shape of any bruises, lacerations, burns, bites, or other skin injuries should be documented in a medical chart as well as with high-quality 35-mm or digital photographs. Inspection for injuries should be thorough and involve all aspects of the neck and head; mouth; extremities, including feet and hands; genitals; anus; buttocks; torso; and back. Obscure sites for inflicted injuries include the ears, especially the posterior aspects, the neck and angle of the jaw, scalp, and the frenula of the lip and tongue. In contrast to accidental injuries, inflicted injuries tend to occur on surfaces away from bony prominences, such as the neck, head, buttocks, trunk, hands, and upper arms. In 1 patient series, approximately 60% of abused children had injuries on the head, face, or neck. Hematomas of the scalp may be detected through palpation or may be visualized on radiographs. Some deeper bruises may not be readily visible for several hours; areas that are painful to palpate may require further examination in 1 to 2 days, when bruises may become apparent. Measurement of skin injuries may assist in determining the mechanism of injury and/or object used to inflict the injury. For example, a child that is kicked may have a discernable shoe imprint, or a knuckle imprint may be apparent if the child was punched.

Bite marks can yield important forensic information; referral to professionals that can gather such information and maintain a chain of custody is advisable. Bite marks, recent or healed, should be carefully measured and photodocumented when possible; an intercanine distance of more than 2 cm suggests a human adult-sized bite. In some facilities, forensic odontologists are available and may use special examination and photographic techniques to analyze bite marks. Fresh bites should be swabbed with sterile, premoistened cotton-tipped applicators for forensic analysis of potential genetic markers found in saliva.

The age of a bruise cannot be determined accurately. Soft tissue swelling is seen more commonly with recent trauma but can persist for several days. The age and developmental capabilities of the infant or child also determine the frequency of bruising. For example, 1 study of infants and toddlers presenting for health maintenance examinations found that 17.8% of infants starting to “cruise” and 51.9% of ambulatory toddlers had bruises; bruises were observed only 2.2% of the time in infants who were not yet cruising. In addition to accidents, bruising may occur secondary to coagulopathies and vasculitides such as idiopathic thrombocytopenic purpura, vitamin K deficiency, Henoch-Schönlein purpura, hemophilia, or von Willebrand disease.

Burn injuries may be chemical, thermal (including exposure to scalding liquids or hot objects), or electrical. The child’s clothing worn during the burn should be collected and may provide information regarding the cause of the burn. Burns inflicted with hot objects can be difficult to differentiate from accidental mechanisms, because both burns may be patterned. The history, number of burns, and continuity of the burn pattern over curved body surfaces may indicate a greater probability of inflicted trauma. Accidental scalds most commonly involve hot liquids pulled or splashed onto the child’s upper extremities, torso, and or neck and head. Inflicted scalds or forced-immersion burns may be well demarcated in pattern, with few or no splash marks. When evaluating an apparent burn injury, other noninflicted causes to consider include chemical burns of the buttocks with senna-containing laxatives, bullous impetigo, and accidents.

**Cranial Injuries**

Head trauma is the leading cause of child abuse fatalities. When compared with child victims of severe accidents, children with abusive head trauma are more likely to have subdural and subarachnoid hematomas, multiple subdural hematomas of differing ages, more extensive retinal hemorrhages, and associated cutaneous, skeletal, and visceral injuries. The inflicted injuries tend to occur in younger patients. Abusive head trauma tends to result in higher mortality and longer hospital stays than does accidental head trauma. Infants with intracranial injuries frequently have no or nonspecific symptoms, so the absence of neurologic symptoms should not exclude the need for imaging. Careful consideration of symptoms, signs, history, and judicious use of other ancillary tests should guide the clinician in determining the need for imaging.

Skull fractures can occur from accidents or inflicted injury. Studies have indicated that simple linear skull fractures can result from short falls of less than 3 ft and that such fractures are usually associated with scalp bruising or swelling. However, it is unknown how many infants and children sustain skull fractures from simple falls, are asymptomatic, and, therefore, never present for a medical evaluation; hence, the incidence of skull fractures among infants who sustain such falls is likely unknown. Abuse should be suspected when there is a history of minor head trauma such as a short fall in children with multiple, complex, diastatic, or occipital skull fractures. Whenever an infant or child presents with a skull fracture, care should be taken to ensure that there are no other injuries.

Conditions that may be confused with abusive head
trauma include glutaric aciduria type 1 (macrocranium, subdural hematoma, sparse intraretinal and preretinal hemorrhages, frontotemporal atrophy) and hemorrhagic disease of the newborn (including risk factors such as home birth, no vitamin K prophylaxis, or breastfeeding).

A fundoscopic examination for retinal hemorrhages should be considered for any infant or young child who is a suspected victim of physical abuse. Under optimal conditions, an ophthalmologist with pediatric experience should conduct an examination of dilated pupils by using indirect ophthalmoscopy. The ophthalmologist should provide documentation of the retinal hemorrhages by photography or detailed annotated drawings. Location, depth, and extent of retinal hemorrhages may distinguish between abusive and nonabusive causes of head trauma.53,54 Retinal hemorrhages occur in approximately 85% of infants and children who are subjected to abusive, repetitive, acceleration-deceleration (shaking) forces with or without impact.48 Although newborn infants may have retinal hemorrhages in the superficial nerve fiber layers, most resolve by 2 weeks of age, and most intraretinal hemorrhages resolve by 4 to 6 weeks of age.49

Thoracoabdominal Injuries
Inflicted injuries that involve the heart are rare and severe. Rib fractures in infants are usually caused by forceful squeezing of the chest50; posterior or lateral rib fractures or multiple rib fractures are especially predictive of abusive trauma.51 Cardiopulmonary resuscitation, whether performed by experienced or inexperienced individuals, is an unlikely cause of rib fractures52 or retinal hemorrhages. Acute rib fractures may be associated with shallow breathing attributable to pain and splinting; in severe cases, a fractured rib may puncture the lung. Alterations in respiratory patterns may also signal central nervous system damage or response to pain. Other rare injuries associated with abusive blows or compressive forces to the chest include hemopericardium, cardiac contusions occurring as a result of abusive blows to the chest, and shearing of the thoracic duct resulting in chylothorax.53,54

Auscultation, performed before palpation, may reveal decreased or no bowel sounds if the child has sustained intraabdominal injury. If the intestines, liver, or spleen have been ruptured, guarding or abdominal muscle rigidity may be noted on palpation. Abdominal bruising is often not seen, even with severe blows to the abdomen.55 In 1 study,56 solid organ injuries were most common in children with accidental and inflicted abdominal trauma, but abused children were more likely to have a hollow viscus injury or both hollow viscus and solid organ injuries than were children with accidental abdominal injuries. In comparison with children who sustain accidental trauma to the abdomen, victims of inflicted intraabdominal injury tend to be younger, are more likely to have delayed presentations to a clinical setting, have a higher mortality rate, and are more likely to have an injury to hollow viscera.5 Liver and pancreatic enzyme tests are helpful in screening children for abdominal trauma, especially when the child presents with acute symptoms or shortly after the incident has occurred. A urinalysis may also lead to the discovery of unexpected trauma to the urinary tract and kidneys. Radiographic studies, including computed tomography, are helpful in determining the types and severity of intraabdominal trauma and are warranted in most cases when the physical examination is unreliable because of patient age, presence of other injuries that may obfuscate the abdominal examination, or the presence of head injury.

Skeletal Injuries
Careful palpation of the legs, arms, feet, hands, ribs, and head may reveal acute or healing (callus formation) fractures. If a fracture is suspected, surfaces should be carefully examined for “grab marks” that may indicate restraint or areas that were pulled or twisted to create the fracture; however, absence of such bruising does not exclude abusive mechanisms of injury. Soft tissue swelling, with or without bruising, may indicate more recent trauma. Many fractures, including rib and metaphyseal fractures, may not be clinically detectable, so a negative clinical examination should not preclude the need for a skeletal radiologic survey when inflicted trauma is suspected, particularly in children younger than 2 years.

Long-bone fractures that should be evaluated carefully for nonaccidental causes include metaphyseal fractures and spiral/oblique fractures, especially in nonambulatory infants; both types of fractures have been associated with accidental mechanisms of injury as well. Accidental causes of lower-extremity spiral or oblique fractures have been described among infants in “exersaucers”57 and in the tibia of newly ambulatory toddlers.58 Osteogenesis imperfecta is a rare congenital disorder that typically presents with bone fragility. Other associated findings are common and include deep-blue sclera, ligamentous laxity, osteopenia, wormian skull bones, denticinogenesis imperfecta, positive family history, and hearing loss. Less common types of this disease may present with fewer and less-severe clinical symptoms.59 Patients with osteogenesis imperfecta are often suspected as victims of abuse before diagnosis, because the history of the injury insufficiently explains the severity of the fracture, and osteopenia may be lacking in occult cases of this disease.60

A complete neurologic assessment, including reflexes, cranial nerves, sensorium, gross motor, and fine motor abilities, should be conducted. Abnormalities may reflect current or past injuries to the central nervous system. Abused children may also have developmental disabili-
ties because of deprivation in the home environment or other causes.

**DIAGNOSTIC TESTING AND CONSULTATIONS**

When abuse is suspected as the cause of an injury, the clinician may conduct tests to screen for other injuries or underlying medical causes for the injury. The extent of diagnostic testing depends on several factors including the severity of the injury, the type of injury, the age of the child, and examination findings. In general, the more severe the injury and younger the child, the more extensive is the need for diagnostic testing for other injuries. Table 1 is a summary of tests, some of which may be used during a medical assessment for suspected abuse.

When 1 child is identified as a suspected victim of abuse, siblings and other child contacts of the suspected abuser should also be assessed for injuries. The extent of the assessment depends on the child’s age, symptoms, and signs; infants and toddlers may require more extensive testing, because symptoms and signs may be less useful in determining the presence of occult inflicted injuries.

**DOCUMENTATION AND DIAGNOSTIC CONSIDERATIONS**

Complete documentation of visible injuries on body diagrams and with photographs is strongly urged and facilitates peer review as well as court testimony, when required. In some regions, investigators from law enforcement or child protective services are specially trained to take forensic photographs. Diagnostic impressions should address whether the explanation adequately correlates with the severity, age, pattern, and distribution of the injury or injuries and the likelihood of nonaccidental causes for the injury. If a child has sustained a serious injury because he or she was left unsu-

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Type of Injury or Condition</th>
<th>Diagnostic Tests</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fractures</td>
<td>Skeletal survey: humeri, forearms, femurs, lower legs, hands, feet, skull, cervical spine, thorax (including oblique views) and lumbar spine, pelvis</td>
<td>1. Recommended for all children with fractures and children with any suspicious injuries under age 2 2. Repeat skeletal survey in 2 wk for high-risk cases 3. Single whole-body films are unacceptable</td>
</tr>
<tr>
<td>Bruises</td>
<td>Tests for hematologic disorders: CBC count, platelets, prothrombin time, partial thromboplastin time, INR, bleeding time; additional testing (eg, factor levels) may be indicated after initial screening tests</td>
<td>1. Recommended when bleeding disorder is a concern because of clinical presentation or family history 2. A DIC screen should be performed for patients with intracranial injury, because intraparenchymal damage can alter coagulation 3. PFa-100: platelet function activity is preferable to bleeding time for establishing platelet function but is not widely available</td>
</tr>
<tr>
<td>Liver injury</td>
<td>Liver enzyme tests: aspartate aminotransferase and alanine aminotransferase</td>
<td>1. May be helpful in diagnosing occult hepatic injury</td>
</tr>
<tr>
<td>Urinary system/renal injury</td>
<td>Urinalysis</td>
<td>1. When used in conjunction with radiographs, may enhance detection of skull fractures 2. Screen for glutaric aciduria type 1 3. IV contrast should be used and is preferable to PO</td>
</tr>
<tr>
<td>Intracranial and extracranial injury</td>
<td>MRI: head/neck</td>
<td>1. May be used for acute rib fractures and subtle, nondisplaced long-bone fractures</td>
</tr>
<tr>
<td>Intracranial and extracranial injury</td>
<td>CT scan: head</td>
<td>1. May be used for acute rib fractures and subtle, nondisplaced long-bone fractures</td>
</tr>
<tr>
<td>Intracranial injury</td>
<td>Urine: organic acids</td>
<td>1. More sensitive than plain radiographs and CT for detecting cervical spine fractures/injury</td>
</tr>
<tr>
<td>Intra-abdominal injuries</td>
<td>Cardiac enzymes: troponin and creatine kinase with muscle and brain subunits (CK-MB)</td>
<td>1. More sensitive than plain radiographs and CT for detecting cervical spine fractures/injury</td>
</tr>
<tr>
<td>Cardiac injury</td>
<td>Radionuclide bone scan</td>
<td>1. Better for acute rib fractures and subtle, nondisplaced long-bone fractures</td>
</tr>
<tr>
<td>Skeletal</td>
<td>Bone-mineralization disorders: rickets</td>
<td>Calcium, alkaline phosphatase, phosphorus, vitamin D, and parathyroid hormone</td>
</tr>
</tbody>
</table>

Tests should be ordered judiciously and in consultation with the appropriate genetics, hematology, radiology, and child abuse specialists. Careful consideration of the patient’s history, age, and clinical findings should guide selection of the appropriate tests. CBC indicates complete blood cell; INR, international normalized ratio; DIC, disseminated intravascular coagulation; CT, computed tomography; IV, intravenous; PO, oral; CK-MB, creatine kinase MB band.

* CT scanning may provide clinically relevant information more expeditiously than MRI in some facilities.
pervised in a dangerous environment, the physician should report suspected neglect or inappropriate adult supervision, including injuries sustained while under the care of an intoxicated adult, to child protective services. When the child is evaluated or tested for other nonabusive causes, documentation should reflect the results of this assessment as well. In general, concern for abuse is greatest for infants younger than 12 months regardless of the severity of the injury.

TREATMENT
Once medical assessment and stabilization are achieved and a referral has been made to investigative agencies, the physician should ensure that the child receives the necessary follow-up services. The child’s primary care physician should be notified, and child protective services should ensure that the family complies with the plan of care. These services should not only include referral to appropriate medical providers but also address the psychological effects of abuse or neglect on the young child, the siblings, and the nonoffending caregiver. Because adult-partner violence commonly co-occurs with child abuse, several family members may require medical and mental health assistance. Medical passports, which are abbreviated medical chart forms usually kept by foster parents and presented at each medical visit, are recommended to optimize treatment regimens in children who are shifted among agencies and individuals during the course of the child abuse investigation.

LEGAL ISSUES
All 50 states have statutes that mandate reporting of suspected child abuse and neglect; the physician is not required to prove abuse before reporting. Familiarity with state laws will ensure that physicians report to the appropriate agency within the required time frame; some states have provided the option of making such a report through the Internet. Information on specific state laws are provided by the Children’s Bureau (Administration for Children and Families, US Department of Health and Human Services; see www.childwelfare.gov/systemwide/laws_policies/search/index.cfm). Many states have laws that permit physicians to evaluate children who are suspected victims of abuse, to conduct tests, and to take photographs without parental consent.

The physician may be required to write a sworn statement of his or her findings and to testify in civil or criminal trial proceedings. Civil hearings include testimony about the safety of the child and the need for appropriate placement with caregivers or state agencies. Judgments are based on a “preponderance of the evidence” with respect to the likelihood of abuse. Criminal hearings involve testimony about the guilt or innocence of an individual with respect to causing the injuries in a child. The burden of proof is greater than that of civil hearings; cases must be proven “beyond a reasonable doubt.” Physicians are expected to testify to the facts on the basis of their knowledge and experience in pediatrics and, when appropriate, in child abuse. As such, they may be asked to render opinions regarding the normal developmental capabilities of children at certain ages as well as the mechanisms of injury, severity of the injury, and prognosis. Physicians should not testify to anything that is beyond their level of knowledge or expertise. Physicians act primarily as scientists and educators in legal settings rather than as child advocates.

CONCLUSIONS
Child physical abuse is a common problem of childhood. The physician must be able to recognize suspicious injuries, conduct a comprehensive and careful examination with appropriate auxiliary tests, critically assess the explanation provided for the injury or injuries, and establish the probability that the explanation does or does not correlate with the pattern, severity, and/or age of the injury or injuries. The physician is responsible for reporting suspected abuse, documenting his or her opinions clearly, and providing the necessary information and expertise to investigative and legal personnel and parents, when appropriate. In addition, pediatricians are uniquely qualified to work with parents and caregivers to prevent abuse by providing anticipatory guidance on normal child behavior and its management. Finally, physicians must advocate that children in foster care who have medical or mental health problems receive the appropriate services and medications and continuity of care through a medical home, and that a medical passport is maintained for these children.

COMMITTEE ON CHILD ABUSE AND NEGLECT, 2006–2007
Carole Jenny, MD, MBA, Chairperson
Cindy W. Christina, MD
Roberta A. Hibbard, MD
Nancy D. Kellogg, MD
Betty S. Spivack, MD
John Stirling, Jr, MD

LIAISONS
David L. Corwin, MD
American Academy of Child and Adolescent Psychiatry
James A. Mercy, PhD
Centers for Disease Control and Prevention

STAFF
Tammy Hurley

REFERENCES
Childhood abuse, neglect, and household dysfunction and the risk of illicit drug use: the Adverse Childhood Experiences Study. Pediatrics. 2003;111:564–572


18. Barr RG, Trent RB, Cross J. Age-related incidence curve of hospitalized shaken baby syndrome cases: convergent evidence for crying as a trigger to shaking. Child Abuse Negl. 2006;30:7–16


Overuse Injuries, Overtraining, and Burnout in Child and Adolescent Athletes

Joel S. Brenner, MD, MPH, and the Council on Sports Medicine and Fitness

ABSTRACT
Overuse is one of the most common etiologic factors that lead to injuries in the pediatric and adolescent athlete. As more children are becoming involved in organized and recreational athletics, the incidence of overuse injuries is increasing. Many children are participating in sports year-round and sometimes on multiple teams simultaneously. This overtraining can lead to burnout, which may have a detrimental effect on the child participating in sports as a lifelong healthy activity. One contributing factor to overtraining may be parental pressure to compete and succeed. The purpose of this clinical report is to assist pediatricians in identifying and counseling at-risk children and their families. This report supports the American Academy of Pediatrics policy statement on intensive training and sport specialization.

INTRODUCTION
Overuse injuries, overtraining, and burnout among child and adolescent athletes are a growing problem in the United States. Although inactivity and obesity are on the rise, the number of children and adolescents who participate in organized or recreational athletics has grown considerably over the past 2 decades. It is estimated that 30 to 45 million youth 6 to 18 years of age participate in some form of athletics. Sports participation is more accessible to all youth, from recreational play and school activities, to highly organized and competitive traveling teams, to pre-Olympic training opportunities. The variety of available, organized sporting activities has also grown from the typical American favorites, such as football, baseball, and soccer, to include lacrosse, field hockey, rugby, cheerleading, and dance, each with its own list of sports medicine concerns. This report will assist the clinician managing young athletes by first defining the medical, psychological, and developmental concerns of intensive, focused athletic participation. In addition, it will highlight specific overtraining issues such as participation in endurance events, weekend athletic tournaments, year-round training on multiple teams, and the multisport athlete. This clinical report should be used in conjunction with the American Academy of Pediatrics policy statement on intensive training and sports specialization in young athletes. There is currently a very small body of scientific evidence pertaining to these issues. Therefore, some of the recommendations are based on committee opinion and/or expertise.
Overuse Injuries
An overuse injury is microtraumatic damage to a bone, muscle, or tendon that has been subjected to repetitive stress without sufficient time to heal or undergo the natural reparative process. Overuse injuries can be classified into 4 stages: (1) pain in the affected area after physical activity; (2) pain during the activity, without restricting performance; (3) pain during the activity that restricts performance; and (4) chronic, unremitting pain even at rest. The incidence of overuse injuries in the young athlete has paralleled the growth of youth participation in sports. Up to 50% of all injuries seen in pediatric sports medicine are related to overuse.

The risks of overuse are more serious in the pediatric/adolescent athlete for several reasons. The growing bones of the young athlete cannot handle as much stress as the mature bones of adults. For example, a young baseball pitcher who has not yet learned proper throwing mechanics (ie, recruiting the entire kinetic chain—from foot to hand—instead of just the arm) is at risk of traction apophysitis of the medial elbow. A young gymnast who performs repetitive hyperextension activities may develop spondylolysis (ie, a stress fracture of the spine), which is an injury particular to the pediatric age group. In addition, young swimmers may not recognize signs of rotator cuff tendonitis, because they may be unable to cognitively connect vague symptoms, such as fatigue or poor performance, as a sign of injury. Identifying youth at risk of overuse injuries is the first step to prevention. Guidelines for parents, coaches, and athletes need to be developed to provide opportunities for education, injury reduction, and early recognition of overuse injuries.

Overtraining
A question often asked of the practitioner who cares for young athletes is, “How much athletic training is too much?” There are no scientifically determined guidelines to help define how much exercise is healthy and beneficial to the young athlete compared with what might be harmful and represent overtraining. However, injuries tend to be more common during peak growth velocity, and some are more likely to occur if underlying biomechanical problems are present.

A sound training regimen is essential, recognizing that although repetition is important, it may induce harm. Sport-specific drills that use a variety of modalities, such as water running for the track athlete on alternate days, may provide similar fitness benefits with less stress to the body. The American Academy of Pediatrics Council on Sports Medicine and Fitness recommends limiting 1 sporting activity to a maximum of 5 days per week with at least 1 day off from any organized physical activity. In addition, athletes should have at least 2 to 3 months off per year from their particular sport during which they can let injuries heal, refresh the mind, and work on strength, conditioning, and proprioception in hopes of reducing injury risk. In addition to overuse injuries, if the body is not given sufficient time to regenerate and refresh, the youth may be at risk of “burnout.”

“Burnout” or Overtraining Syndrome
Burnout, or overtraining syndrome, has been well described in the literature for adult athletes, but little is found regarding its applicability in youth. The overtraining syndrome can be defined as a “series of psychological, physiologic, and hormonal changes that result in decreased sports performance.” Common manifestations may include chronic muscle or joint pain, personality changes, elevated resting heart rate, and decreased sports performance. The pediatric athlete may also have fatigue, lack of enthusiasm about practice or competition, or difficulty with successfully completing usual routines. Burnout should be recognized as a serious sequela of overtraining syndrome. Prevention of burnout should be addressed by encouraging the athlete to become well rounded and well versed in a variety of activities rather than 1 particular sport. The following guidelines are suggested to prevent overtraining/burnout:

1. Keep workouts interesting, with age-appropriate games and training, to keep practice fun.
2. Take time off from organized or structured sports participation 1 to 2 days per week to allow the body to rest or participate in other activities.
3. Permit longer scheduled breaks from training and competition every 2 to 3 months while focusing on other activities and cross-training to prevent loss of skill or level of conditioning.
4. Focus on wellness and teaching athletes to be in tune with their bodies for cues to slow down or alter their training methods.

Endurance Events
Endurance athletic events (triathlons, marathons, and half-marathons) are becoming more popular in the United States, and legitimate concerns have been raised for the safety of youth participating in these events. The American Academy of Pediatrics has stated that triathlons for children and adolescents are reasonably safe as long as the events are modified to be age appropriate. Specifically, such events should be of shorter duration/length, and careful attention should be given to safety and environmental conditions. Children and adolescents must be properly trained to avoid hypothermia or hyperthermia, overtraining, overuse injuries, and burnout.
Recent concerns regarding the participation of children in marathon running has led to different opinions being expressed in the literature. There is, at present, no scientific evidence that supports or refutes the safety of children who participate in marathons. There are no recorded data on injuries sustained by children who run marathons. Marathon training requires a gradual increase in total weekly mileage, which may be less than or equal to the total weekly distance that is generally logged by high school cross-country teams (35–40 miles). Regardless, a clearly devised weekly plan, ensuring that safe running conditions are in place, and the provision of proper education on endurance activities (including environmental conditions and appropriate hydration) should all be part of the training process. A critical environmental safety concern is the ambient temperature and relative humidity, because a child is less able than an adult to handle heat stress. Weather-related guidelines have been set for all marathons, and these guidelines should be strictly enforced by the medical director for all youth endurance events. Ultimately, there is no reason to disallow participation of a young athlete in a properly run marathon as long as the athlete enjoys the activity and is asymptomatic.

Weekend Athletic Tournaments
Weekend-long sports tournaments for soccer, baseball, or tennis are common across the country. Often, these athletes are actively participating at least 6 hours each day in their sport and are exposed to the associated weather elements for an additional 2 to 3 hours. The risks associated with these events include heat-related illness, nutritional deficiencies, overuse injuries (eg, pitching in multiple games over a 48-hour span), and burnout from having a lack of “free time.” Research examining the possibility of fatigue contributing to an increased injury risk in the tournament situation does not exist, but the general overtraining-prevention guidelines outlined earlier should also apply.

Year-Round Training on Multiple Teams
Single-sport, year-round training and competition is becoming more common for children and adolescents. A focus on participating in 1 sport, or single-sport specialization, to improve, advance, and compete at the highest level may drive youth to participate for long hours daily on 1 or more teams at a time. This is common in soccer, baseball, and gymnastics. The motivation behind this overinvolvement may be induced by the child or parent. As more young athletes are becoming professionals at a younger age, there is more pressure to grab a piece of the “professional pie” to obtain a college scholarship, or to make the Olympic team. Most young athletes and their parents fail to realize that, depending on the sport, only 0.2% to 0.5% of high school athletes ever make it to the professional level. Yet, youth continue to specialize in 1 sport while participating on multiple teams and risk overuse and/or burnout if there is no break from athletics during the year. Young athletes who participate in a variety of sports have fewer injuries and play sports longer than those who specialize before puberty.

Multisport Athlete
Well-rounded, multisport athletes have the highest potential to achieve the goal of lifelong fitness and enjoyment of physical activity while avoiding some of the pitfalls of overuse, overtraining, and burnout provided that they participate in moderation and are in tune with their bodies for signs of overuse or fatigue. Many youth will play multiple sports throughout the year either simultaneously or during different seasons. They may do this because they enjoy multiple sports or because their coach or parent pushes them to participate in other sports to condition them for their primary sport or in hopes of being noticed by college or professional scouts. There may be additional pressures from other coaches who wish to better their team by calling on well-rounded athletes from other sports. Multisport athletes are at risk of overuse injuries if they do not get sufficient rest between daily activities or if they do not get a break between seasons. Multisport athletes who participate in 2 or more sports for which the major emphasis is the same body part (eg, swimmers and baseball pitchers) are at higher risk of overuse injuries than are those who participate in sports that have a different emphasis (eg, track and golf).

What Is the Goal of the Athlete?
The ultimate goal of youth participation in sports should be to promote lifelong physical activity, recreation, and skills of healthy competition that can be used in all facets of future endeavors. As providers of care for youth, it is important to obtain a physical activity history (type of activity, frequency, duration) and take the opportunity to promote healthy participation and preventive care measures. Education of parents, athletes, and coaches must be part of the plan to promote fun, skill development, and success for each individual athlete. Skilled young athletes must be mentored carefully to prevent overparticipation, which may affect them physically as well as psychologically. The parent or pediatrician may wonder how hard a child should be pushed to train and compete. Ultimately, it is important for the practitioner to discuss the underlying motivation for sport participation with the athlete, the parent, and, possibly, the coach. Unfortunately, too often the goal is skewed toward adult (parent/coach) goals either implicitly or explicitly. The parent often hopes the child will get a scholarship, become a professional athlete, or fulfill the parents’ unfulfilled childhood dreams. It is best to identify and focus on the child’s motivation and goals to provide guidance.
GUIDANCE FOR THE CLINICIAN

1. Encourage athletes to strive to have at least 1 to 2 days off per week from competitive athletics, sport-specific training, and competitive practice (scrimmage) to allow them to recover both physically and psychologically.

2. Advise athletes that the weekly training time, number of repetitions, or total distance should not increase by more than 10% each week (eg, increase total running mileage by 2 miles if currently running a total of 20 miles per week).

3. Encourage the athlete to take at least 2 to 3 months away from a specific sport during the year.

4. Emphasize that the focus of sports participation should be on fun, skill acquisition, safety, and sportsmanship.

5. Encourage the athlete to participate on only 1 team during a season. If the athlete is also a member of a traveling or select team, then that participation time should be incorporated into the aforementioned guidelines.

6. If the athlete complains of nonspecific muscle or joint problems, fatigue, or poor academic performance, be alert for possible burnout. Questions pertaining to sport motivation may be appropriate.

7. Advocate for the development of a medical advisory board for weekend athletic tournaments to educate athletes about heat or cold illness, overparticipation, associated overuse injuries, and/or burnout.

8. Encourage the development of educational opportunities for athletes, parents, and coaches to provide information about appropriate nutrition and fluids, sport safety, and the avoidance of overtraining to achieve optimal performance and good health.

9. Convey a special caution to parents with younger athletes who participate in multigame tournaments in short periods of time.

COUNCIL ON SPORTS MEDICINE AND FITNESS, 2005–2006
Eric W. Small, MD, Chairperson
David T. Bernhardt, MD
Joel S. Brenner, MD, MPH
Joseph A. Congeni, MD
Jorge E. Gomez, MD
Andrew J. M. Gregory, MD
Douglas B. Gregory, MD
Teri M. McCambridge, MD
Frederick E. Reed, MD
Stephen G. Rice, MD, PhD, MPH

Paul R. Stricker, MD
Bernard A. Griesemer, MD

LIAISONS
Claire M. A. Le Blanc, MD
Canadian Paediatric Society
James Raynor, MS, ATC
National Athletic Trainers Association

STAFF
Jeanne Lindros, MPH
Anjie Emanuel, MPH

REFERENCES
Score for Neonatal Acute Physiology (SNAP) or Vermont Oxford Risk-Adjustment Model for Very Low Birth Weight Infants?

To the Editor.—

We read with great interest the article by Zupancic et al,¹ which compared the revised Score for Neonatal Acute Physiology (SNAP-II) and revised Score for Neonatal Acute Physiology Perinatal Extension (SNAPPE-II) scores with the Vermont Oxford Network risk-adjustment algorithm (VON-RA) in a large cohort of term and preterm infants from North America. In very low birth weight (VLBW) infants, both scores performed equally well, as judged by their receiver operating characteristic (ROC) analysis results, the area being 0.86 for the SNAPPE-II and 0.85 for the VON-RA.

Given that treatment policies of small preterm infants are different in different countries, it is of interest to know how these scores perform in other settings.

We previously published an article² in which we obtained, in VLBW infants, an ROC area of 0.84 for the SNAPPE-II, which is very similar to that found in the Zupancic et al study. In infants without congenital anomalies, the ROC area was 0.86. Here we report the performance of the VON-RA model in this data set: using the original logistic coefficients (kindly provided by J. Horbar, MD, and J. Carpenter, MS), the VON-RA model had an ROC area of 0.906 for infants without congenital anomalies. This value was confirmed in the complete cohort of 2070 VLBW infants³ admitted to 14 NICUs in Lombardy, northern Italy, where ROC areas of 0.906 and 0.907 for infants with and without congenital anomalies, respectively, were found.

We conclude that in our setting, the SNAPPE-II had a performance similar to that found in the Zupancic et al study (and in previous articles), whereas the VON-RA model had better discrimination. Interestingly, different centers showed greater variability in SNAPPE scores (mean score: 18–36 [P < .0001] by analysis of variance) than in VON-RA scores (P = .22).

We agree with the authors¹ that a measure of severity of illness that takes into account individual characteristics of the infant (eg, urine output, core temperature, and pH, as in the SNAPPE-II) should be better, theoretically, than a risk score based only on fixed covariates (eg, gender, multiple pregnancy, outborn status, etc). Unfortunately, as the authors admit, no severity scoring system can be used to judge an individual infant’s risk of mortality because of the wide confidence limits of the estimate. In fact, an ROC area represents the probability that the score of a randomly selected infant who died will be greater than the score in of a randomly selected normal (surviving) infant.⁴ Thus, even an ROC area of 0.86 would not look so impressive, meaning a 14% overlapping of scores between deaths and survivals, which is a figure that is too high to guide therapeutic decisions.

On the other hand, we found that for risk adjustment in groups of subjects (eg, for comparing hospital performances), the VON-RA worked very well and can be calculated from data collected routinely, as opposed to the SNAPPE, which requires careful and time-consuming collection of data for this purpose.

Finally, congenital anomalies add an important risk of death, even without increasing physiologic instability (ie, even without increasing the SNAPPE-II), and up to now no clinically useful way to quantify this risk has been found. We think that an empirical classification of congenital anomalies, calculated from the very large database of the VON, is a major improvement, and we hope that this classification will be made publicly available soon.

Luigi Gagliardi, MD
Division of Neonatology and Pediatrics,
Ospedale Versilia,
I-55043 Lido di Camaiore, Italy
Italian Neonatal Network
Asthma as a Barrier to Children’s Physical Activity

To the Editor.—

We read with interest the recent report from Glazebrook et al.1 They found a high prevalence of obesity (21.4%) in their population of outpatients with asthma and that asthma acted as a barrier to exercise in these children. We performed a retrospective review of a pulmonary-function database maintained at our institution (1988–2005), which we think lends an interesting additional perspective to the findings of Glazebrook et al. Within this database, 1095 children with a diagnosis of asthma were referred for testing. The BMI was calculated and categorized as normal (5th–85th percentile), overweight (85th–95th percentile), or obese (>95th percentile). We found that in this population with a median age of 10.28 years (range: 4.4–25 years), the body composition was normal in 691 patients (63%), overweight in 178 (16%), and obese in 226 (21%). Lung-function testing revealed that the ratio of forced expiratory volume in 1 second to forced vital capacity (FEV1/FVC) was slightly lower in the obese group (median: 79% vs 83% in the normal-BMI group; \( P = .001 \)), and functional residual capacity (FRC) was significantly lower in the obese group (median: 93% predicted vs 103% predicted in the normal-BMI group; \( P < .001 \)). There were no differences between BMI groups in the degree of improvement in specific airway conductance after administration of inhaled bronchodilator.

These findings and those of Glazebrook et al raise an important question: are children with asthma less likely to exercise and thus be at high risk of obesity (because they are limited by their asthma symptoms), or does obesity itself cause asthma to be more difficult to control? Or, are both true? Similar to Glazebrook et al, we found a high prevalence of obesity in patients with linked to interventions that show actual practice or outcomes change.

John A. F. Zupancic, MD, ScD
Division of Newborn Medicine
Harvard Medical School
Boston, MA 02115

Shoo K. Lee, MD, PhD
Integrated Centre for Care Advancement Through Research Edmonton (iCare)
University of Alberta
Edmonton, Alberta, Canada T6G 2M8

Gabriel J. Escobar, MD
Perinatal Research Unit
Division of Research
Kaiser Permanente Medical Care Program
Oakland, CA 94612
doi:10.1542/peds.2007-0671
asthma who were referred for lung-function testing. However, asthma is also more prevalent in overweight children. The high prevalence of asthma in overweight and obese children has been attributed to abnormal baseline lung mechanics or to inflammatory mediators produced by adipose tissue. However, despite earlier studies that showed airway hyperresponsiveness to exercise and methacholine in obese subjects with asthma, we did not find increased airway responsiveness to bronchodilators in obese children compared with those in other BMI groups.

A relative reduction in bronchodilator responsiveness compared with bronchial hyperreactivity to constricting stimuli in the obese patient with asthma could be a result of airway remodeling (not highly likely at this age), or it may be related to airway/parenchymal/chest-wall interactions such as the effect of obesity on slowly cycling latch bridges in airway smooth muscle. According to this hypothesis, the ability of airway smooth muscle to relax when exposed to bronchodilators depends on its ability to break and reattach actin-mysin links. Low FRC and tidal-volume states such as obesity, by reducing tidal stretch on airway smooth muscle, reduce the probability of actin-mysin detachment/reattachment, in effect “freezing” the attachments in a “latch state.” This could explain less-than-expected bronchodilator responsiveness in obese subjects with asthma. In addition, the airway “compressive” effects of a low FRC may limit the amount that the airways can dilate in an obese subject with asthma.

What are the implications of these findings for the cause-and-effect question posed above? Obese subjects with asthma have asthma symptoms that are more difficult to control. Increased bronchial lability to exercise unaccompanied by increased bronchodilator responsiveness could likewise make exercise-induced asthma more difficult to control in the obese subject, which leads to less willingness to exercise and, in turn, increases the risk of obesity. We agree with Glazebrook et al that barriers to exercise in obese children with asthma need to be addressed. It would be of interest to study whether intermittent increases in tidal volume in obese subjects, such as those engendered by incentive spirometry, could help increase bronchodilator responsiveness by “unfreezing” the latch state. Our findings suggest that because there is increased bronchial lability with exercise that is not matched by increased bronchodilator responsiveness in obese subjects, more importance needs to be placed on preventative therapy, including appropriate use of antiinflammatory controller medications and control of obesity itself.

Daniel J. Weiner, MD  
Division of Pulmonary Medicine  
Children’s Hospital of Pittsburgh  
Pittsburgh, PA 15213

REFERENCES

10. Saint-Pierre P, Bourdin A, Clanez P, Daures JP, Godard P. Are overweight asthmatics more difficult to control? Allergy. 2006;61:79–84

In Reply.—

We welcome the interest shown by Weiner and McDonough in our article on asthma as a barrier to physical activity. In our study, we used the International Obesity Task Force definition of overweight and obesity, whereas Weiner and McDonough used the 85th and 95th percentile for BMI for overweight and obesity, respectively. However, we were interested to read that they found the same prevalence of obesity in children with asthma as we did in our study (21%).

They raise the question of whether asthma leads to inactivity and obesity or whether obesity itself causes asthma. Prospective studies (which can establish whether obesity precedes asthma) have shown inconsistent results, with some studies showing no relationship and others showing an effect in either boys or girls but not both genders. Whichever is the case, the children in our study, and their parents, perceived asthma as a barrier to exercise. The implications of this is that they
will be less likely to exercise, which could be an important tool for them in tackling obesity. If obesity itself has a causative role in asthma in these patients, then exercise and weight loss may be accompanied by an improvement in asthma symptoms as well as the well-known health benefits and improvement in self-esteem.

We are currently undertaking a pilot study to examine a structured exercise program for children with asthma. We will evaluate the young people’s attitude to exercise before and after they have completed the program and that of their parents. If the program proves to be practical and acceptable in this pilot study, we will progress to a randomized, controlled trial of a structured program versus written advice.

Whether obesity causes asthma or is simply the consequence of a sedentary lifestyle induced by a fear of asthma symptoms, exercise offers an intervention that can improve self-esteem without loss of symptom control.

Alan Smyth, MA, MB, BS, MRCP, MD, FRCPCCH
Division of Child Health
Ian A. Macdonald, PhD
School of Biomedical Sciences
Amy C. McPherson, PhD, CPsych
School of Nursing
Judy A. Swift, PhD
Division of School of Biosciences
University of Nottingham
Nottingham NG7 2RD, United Kingdom
Cristine Glazebrook, RGN, PhD, CPsych
School of Community Health Sciences
Queens Medical Centre
University of Nottingham
Nottingham NG7 2UH, United Kingdom

REFERENCES

Subarachnoid Hemorrhage in a Young Child With Sickle Cell Disease: Is Transcranial Doppler Helpful?

To the Editor.—

We read with interest the article by Strouse et al.1 This work provided an excellent review of intracranial hemorrhage in children with sickle cell disease (SCD). The study also identified risk factors for the development of primary hemorrhagic stroke that included hypertension, red blood cell transfusion, and corticosteroid therapy. We would like to report a case of atraumatic subarachnoid hemorrhage (SAH) with intriguing transcranial Doppler (TCD) results in a child with SCD that may add to the risk profile described in the Strouse et al article.

An 8-year-old girl with hemoglobin SS disease was admitted to Children’s Hospitals and Clinics of Minnesota for management of a vasoocclusive painful event. On the second hospital day, she developed acute chest syndrome. She received 2 red blood cell transfusions as part of her management. She did not receive corticosteroid therapy. By hospital day 8 she was prepared for discharge but developed headache, hypertension, and a generalized seizure. A computed tomography scan, MRI, and lumbar puncture confirmed the presence of a left frontal SAH. Interestingly, the patient had 2 TCD examinations performed during her hospitalization, one on the day of her seizure and one 6 days earlier.

The patient’s baseline hemoglobin level was 7.4 g/dL and decreased to 6.3 g/dL on hospital day 2. After red blood cell transfusion, her hemoglobin level increased by 65% to 10.4 g/dL. Steady-state blood pressures were normal at 117/61 mm Hg. On the day of her SAH she had 2 episodes of hypertension with readings of 140/91 and 137/97 mm Hg. A mean decrease in velocity of 91 cm/second was seen on the TCD that was performed 6 hours before the patient’s seizure when compared with the study performed 6 days earlier. Results are shown in Table 1.

Isolated SAH is quite rare in children with SCD. Our patient represents only the seventh reported case in a child under the age of 9 years.1–3 She had exhibited 2 of the reported risk factors for primary hemorrhagic stroke: red blood cell transfusion and hypertension. Results of a magnetic resonance angiogram and conventional cerebral angiography were normal. She recovered fully and continues to do well without neurologic sequelae. Of particular interest in this case is the TCD examination on

<table>
<thead>
<tr>
<th>Table 1 TCD Results</th>
<th>Left, cm/sec</th>
<th>Right, cm/sec</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Day 2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>204</td>
<td>204</td>
</tr>
<tr>
<td>MCA</td>
<td>180</td>
<td>222</td>
</tr>
<tr>
<td>BF</td>
<td>154</td>
<td>156</td>
</tr>
<tr>
<td>ACA</td>
<td>164</td>
<td>169</td>
</tr>
<tr>
<td>DICA</td>
<td>186</td>
<td>150</td>
</tr>
<tr>
<td>PCA</td>
<td>117</td>
<td>154</td>
</tr>
<tr>
<td><strong>Day 8</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M1</td>
<td>65</td>
<td>77</td>
</tr>
<tr>
<td>MCA</td>
<td>90</td>
<td>97</td>
</tr>
<tr>
<td>BF</td>
<td>81</td>
<td>88</td>
</tr>
<tr>
<td>ACA</td>
<td>85</td>
<td>86</td>
</tr>
<tr>
<td>DICA</td>
<td>74</td>
<td>99</td>
</tr>
<tr>
<td>PCA</td>
<td>71</td>
<td>61</td>
</tr>
</tbody>
</table>

M1 indicates M1 segment of middle cerebral artery; MCA, middle cerebral artery; BF, bifurcation vessel; ACA, anterior cerebral artery; DICA, distal internal carotid artery; PCA, posterior cerebral artery.
the afternoon of the hemorrhage. Low velocities or rapid changes in velocities on TCD may be useful in identifying vasospasm associated with SAH in adults.6,7 Perhaps vasospasm played a role in our patient’s hemorrhage. Of note, her TCD velocities on hospital day 2 may have been affected by her acute painful event, lower hemoglobin level, and narcotic use.

The role of TCD in the setting of pediatric SAH is unclear. With additional study, TCD findings could potentially be included in the risk assessment for primary hemorrhagic stroke in children with SCD.

Christina S. Ross
Jane M. Hennessy, RN, MPH, CNP
Stephen C. Nelson, MD
Pediatric Hematology/Oncology
Children’s Hospitals and Clinics of Minnesota
Minneapolis, MN 55404

REFERENCES

doi:10.1542/peds.2007-0198

In Reply.—

The observation of Ross et al provides a serendipitous insight into the possible effects of transfusion and subarachnoid hemorrhage (SAH) on cerebral blood flow velocities (CBFVs) in children with sickle cell disease (SCD). Their patient, an 8-year-old girl with hemoglobin SS, had abnormally high transcranial Doppler (TCD) velocities of the bilateral middle cerebral arteries (204 cm/second on the left and 220 cm/second on the right) before blood transfusion for acute chest syndrome. CBFV is inversely proportional to hematocrit,1 so worsening anemia may have been partially responsible for the elevation. Increases in metabolic requirements of the brain, such as with a fever or hypoxia, also increase CBFV as the brain increases blood flow to meet metabolic demands.2 However, although not seen on angiography, the degree of elevation suggests that other processes, such as fixed stenosis or vasoconstriction, may have also contributed. Elevated velocity was associated with hemorrhagic stroke in 9 children screened for or participating in the Stroke Prevention Trial in Sickle Cell Anemia (STOP) study, but this association was not as strong as that seen for ischemic stroke.3 Because anemia and other factors are important contributors to increases in CBFV, TCD has only been validated as a screening study for stroke when the patient is not acutely ill. This is the most widely accepted indication for TCD in children, but it is also used to evaluate for cardiac or pulmonary shunts, to document absent cerebral blood flow in brain death, and to evaluate for vasospasm after SAH.4

The patient had a profound global decrease in CBFV at least 6 hours before her seizure. The decrease in velocity was likely multifactorial, with some attributable to the 65% increase in hematocrit and the decrease in the percentage of hemoglobin S4 and the rest possibly attributable to the SAH. Acutely, CBFV is decreased globally in SAH, with the greatest decreases in the most severe hemorrhages.6 Subsequently, CBFV increases, typically between 4 and 12 days after the initial hemorrhage, and is usually secondary to vasospasm or proliferative arteriopathy.7 Existing cutoffs for vasospasm (>120 cm/second for mild and >180 cm/second for severe) have only been validated for adults without significant anemia.

Of the general population with spontaneous SAH, 15% to 20% will not have an etiology identified.6 The exact mechanism of SAH in children with SCD is unknown. It may result from an aneurysm or arteriovenous malformation, as in other children. Patients with preexisting cerebrovascular disease may have bleeding from the leptomeningeal collateral vessels that develop in patients with multiple distal cerebral vessel branch occlusions.8 However, this child had an evaluation with magnetic resonance and conventional angiography, which did not demonstrate any evidence of aneurysm, obstructive vasculopathy, or previous ischemic stroke. As we discussed in our article, SAH in children with SCD commonly occurs after recent blood transfusion, as in this child’s case.9 This may be a result of decreased ability to autoregulate cerebral blood flow after blood transfusion, perhaps related to changes in blood volume, the effects of vasoactive substances in stored blood, or other causes.

With no etiology identified for the SAH, Ross et al face difficult decisions in treating their patient. There are no firm data regarding acute management of SAH in children with SCD, but the management tools applied to patients with SAH but without SCD seem reasonable: stabilization in a neurologic or pediatric ICU, consideration of nimodipine to decrease the risk of vasospasm, and a search for and occlusion of potential aneurysm or arteriovenous malformation. Most hematologists would recommend exchange blood transfusion to decrease the
hemoglobin S level to <30% to maximize tissue oxygenation during the acute period of recovery from the SAH. Existing data are inadequate to guide long-term therapy aimed at reducing the risk of recurrent neurologic events. Certainly the patient’s CBFV should be ascertained again once she has fully recovered from the acute event; if her CBFV is still elevated, a chronic blood-transfusion regimen would be recommended on the basis of the results of the STOP study. However, in patients without chronically elevated CBFV or history of ischemic stroke, the optimal management is not known. Usual care or chronic blood-transfusion therapy has been used in children with SCD and SAH. Ultimately, the decision on treatment approach must then be based on professional judgment and discussion with the patient and family. We appreciate the authors’ thoughtful comments and concur that the use of TCD in children with SAH requires additional study.

John J. Strouse, MD
Division of Pediatric Hematology
Department of Pediatrics
Johns Hopkins University School of Medicine
Baltimore, MD 21205

Monica L. Hulbert, MD
Division of Pediatric Hematology/Oncology
Riley Hospital for Children
Indianapolis, IN 46202

Michael R. DeBaun, MD, MPH
Department of Pediatrics
Washington University School of Medicine
St Louis, MO 63110

Lori C. Jordan, MD
Department of Neurology
James F. Casella, MD
Division of Pediatric Hematology
Department of Pediatrics
Johns Hopkins University School of Medicine
Baltimore, MD 21205

REFERENCES

Neonatal Brain Volumetric Studies: Regression Analysis and Interpretation

To the Editor.—

Volumetric techniques are increasingly used to investigate the impact of injury or intervention on the neonatal brain. Parikh et al recently reported an interesting article that postnatal dexamethasone therapy is followed by reduced cerebral tissue volumes. In view of relatively poor MRI signal contrast between gray and white matter in neonates, researchers overcome challenges in regard to image acquisition and segmentation techniques. However, there remains a big challenge in how best to analyze and interpret the volumes measured. When interpreting the results of neonatal brain volumetric studies, there is often a need to address possible confounding factors. The decision on what confounding factors/covariates to include in the regression model is critical for the conclusions of the study. Equally critical is the way that these possible confounding factors are being tested for and analyzed in the regression model.

With regard to the hypothesis tested, previous scientific evidence is usually a good way to start finding which possible covariates to test for. In a landmark article in 1998, Hüppi et al addressed clinical parameters that correlate with neonatal cerebral volumes. Moreover, depending on the selection criteria and hypotheses, different samples may display different significant covariates. The statistical significance of group differences on demographic and clinical characteristics may be of some help but is limited. Factors not significantly different between groups may well be proven statistically significant when entered in the regression model to test the primary hypothesis, and vice versa. Also, the influence of a covari-
ate on the primary outcome measure may partially over-
lap with the influence of a different covariate, which makes the combination tested crucial.

In regard to the regression model, it is a good explor-
atory approach to enter and test the possible covariates initially 1 by 1. This can be followed by a stepwise regression that involves the more statistically significant ones. It is safer to use larger \( P \) values (eg, .1 or .2) for entry criteria in the stepwise regression than for exit (eg, .05 or .1). The covariates that have the most influence on the primary outcome (cumulative effect if >1) should stay in the model, and then the primary outcome measure-
cments can be adjusted accordingly. It is important to
note whether comparisons for the primary outcome were statistically significant before and after the regression
analysis.

With respect to the above and in terms of cortical volume, along with postmenstrual age at scan, scaling effects and size differences between subjects and groups at scan could have an influence on neonatal cerebral volumes and be a significant covariate. Also, uncomplicated germinal matrix-intraventricular hemorrhage could be a significant covariate for the cortical volume. For the reader, understanding is benefited from ade-
quate details on the demographic and clinical character-
istics of the subjects. The presentation of the raw data and their statistics, before adjustments, is also essential for the interpretation of neonatal brain volume studies.

George T. Vasileiadis, MD, MSc
School of Human Development
University of Nottingham
Nottingham NG7 2UH, United Kingdom
London, Ontario, Canada N6A 4V2

REFERENCES
1. Parikh NA, Lasky RE, Kennedy KA, et al. Postnatal dexametha-
2. Hüppi PS, Warfield S, Kikinis R, et al. Quantitative magnetic resonance imaging of brain development in premature and ma-
3. Vasileiadis GT, Gelman N, Han VKM, et al. Uncomplicated in-
traventricular hemorrhage is followed by reduced cortical vol-
ume at near-term age. Pediatrics. 2004;114(3). Available at: www.pediatrics.org/cgi/content/full/114/3/e367
doi:10.1542/peds.2007-0364

In Reply.—

We thank Dr Vasileiadis for his letter highlighting some key aspects of our article and observational data sets in general.1 We agree with him about the chal-

lenges of assessment and adjustment for confounding factors to uncover any true associations between risk factors of interest and cerebral volumes. His letter de-
scribes one of several statistically acceptable approaches

for variable selection and model building.2–4 As outlined in the statistical-analysis section of our article, we used prespecified criteria for inclusion of covariates to avoid bias in selecting the regression model. The goal of vari-

able selection is to achieve a balance between simplicity (with as few independent variables as possible) and fit (with as many independent variables as needed). Al-

though the inclusion of additional covariates to improve model fit seems desirable, adjustments made by models that include too many covariates, even when statistically significant, may be spurious. Especially for preliminary smaller studies such as ours, inclusion of too many vari-

ables in the final model will produce numerically unsta-

ble models because of overfitting.5 Therefore, a simpler model may actually estimate the true relationship more accurately than a model with many independent vari-

ables.

It is evident from the work of Hüppi et al,6 and the strong association we reported, that postmenstrual age at MRI scan has the greatest impact on brain volumes. As such, we adjusted all of our analyses for this important covariate. Including birth weight in the final model ad-
dressed group differences in size at baseline. Adjusting brain volumes for differences in body size at MRI scan, as suggested by Dr Vasileiadis, could understated the true effect of steroids on regional and total brain volumes if the neurologic effects of dexamethasone are mediated through adverse effects on growth/size because of the reduction in growth resulting from dexamethasone. A more common approach in assessing regional volumes is to adjust for total intracranial volume as a covariate in the regression equation. Because dexamethasone may reduce overall brain growth and intracranial volume, this approach could also understate regional volume differences. In any case, adjusting for total intracranial volume resulted in little or no change in the \( P \) values in our 5 regions of interest (\( P \) values ranged from .07 to .001).

With respect to including uncomplicated germinal matrix-intraventricular hemorrhage as a covariate, we did not find an association between cerebral volumes and this potential determinant or, as we tested in our study, white matter injury, a more powerful determin-

ant of brain volumes.7 As stated in our discussion, our significant associations may have resulted, in part, from incomplete adjustment of measured and/or unmeasured confounders. Such a problem, of course, is not unique to this area of investigation; any observational study may suffer from residual confounding. Confounding is less likely to occur in randomized trials, which is the best study design for resolving the effects of postnatal steroids on brain development in high-risk preterm infants. We are currently in the process of completing such a trial8 and hope to address these lingering questions.
REFERENCES

doi:10.1542/peds.2007-0725

Early Autism Identification

To the Editor.—

Applying an autism-specific screening tool at 18 months has been recommended by experts. In a recent commentary, repeat autism screening at 24 months, after the initial 18-month administration, is also advised by experts to detect autism concerns. These recommendations are currently evidence based and appropriate to known early social communication development. The proposed screening approach, combined with critically needed understanding by primary pediatric providers, has significant potential for success.

However, certain critical issues deserve attention. Autism-specific interventions and programs need to be accessible and provided at these age levels to support appropriate public health screening measures. Lack of specific appropriate interventions would weaken efficacy of any screening program and, specifically, efficacy of early intervention for autism-affected children.

Undue parent anxiety and the absence of appropriate intervention are unacceptable for any screening program. Realistic public health issues of poor health care access and lack of appropriate available autism-specific interventions to children at the ages being screened are serious potential limiting factors. Appropriately addressing these pressing issues is urged from both humanitarian care and public health–screening principles.

Efforts to promote optimal pertinent primary provider education and understanding of typical and disordered social communication are also essential.

Furthermore, in view of current Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV), criteria that capture autistic-disorder presentation by 36 months, repeat administration of an autism-specific screen should arguably be considered by or at 36 months of age. Although screening at 18 and 24 months is preferable and expected to attain longer early intervention, identifying most children of concern, screening again at 36 months has the potential to identify some children with later-presenting concerns. Supporting consideration for additional later screening is the understanding that screening tools provide <100% detection, that some children may present later as encompassed by DSM-IV criteria, and that current reality includes later detection by parents and providers than is ideally desired.

Additional research is needed to address potential success and cost-efficacy of positive identification by screening to further support and guide these screening recommendations.

Penelope J. Pivalizza, MD, MPH, FAAP
Developmental Pediatrics
Baylor College of Medicine
Houston, TX 77030

Developmental Pediatrics
Texas Children’s Hospital
Houston, TX 77030

REFERENCES

doi:10.1542/peds.2007-0509

In Reply.—

All the issues raised by Dr Pivalizza are pertinent to developmental surveillance and screening for autism. We agree that access to health care is pivotal to any
public health effort. The American Academy of Pediatrics (AAP) is committed to every child’s access to a medical home, including children with autism and other developmental disorders.1 The medical home should be linked to an array of specialty services with which children who screen positive for autism can be evaluated to confirm the diagnosis. The child should also be referred to a local early intervention program, coincident with such evaluation.2,3 In the United States, early intervention is now universally available for children with autism through part C of the Individuals With Disabilities Act.3

We share Dr Pivalizza’s concern that local factors may make these services difficult to access or that the provided services may be suboptimal for the child’s particular condition. Work needs to be done to both improve subspecialty access for all medical homes and ensure the availability of good-quality evidence-based interventions for all children who are diagnosed with an autism spectrum disorder. Early intervention should be available through local government-funded services or through health insurance, but such coverage is sometimes less than optimal. The AAP has several resources to help pediatricians and families. The Department of Practice and Quality Improvement (www.aap.org/moc/reimburse/default.cfm) can assist pediatricians who are helping families obtain coverage for such services, including intensive behavioral services.

Provider education is a key to all of these efforts. Through its Expert Panel on Children With Autism, the AAP is developing a tool kit for primary care providers to assist them in the identification and care of children with autism. Developmental surveillance for this developmental disorder is a continuous process that should be performed at all visits. Although the Policy Revision Committee on Developmental Surveillance and Screening and the Expert Panel on Children With Autism have recommended autism-specific screening at 18 and 24 months, the pediatrician should perform it at any age if they have a high index of suspicion for this disorder on the basis of parental concerns or other forms of developmental surveillance.

The cost-effectiveness of autism-specific screening seems self-evident, because it can be performed at a minimal cost (relative value unit: 0.36).4 The benefits of early identification and early intervention of autism are also well established; there exists a substantial body of research that supports the effectiveness of early intensive behavioral interventions for children with autism.5,6

However, we agree with Dr Pivalizza that research needs to be done to determine the cost-effectiveness of developmental and autism-specific screening to “support and guide” our recommendations. Additional study that examines the cost-effectiveness of currently available intervention services or autism-specific services (such as the 25 hours/week recommended for children with autism spectrum disorders6,7) is also necessary for our profession to inform public policy initiatives and to more effectively advocate for the best use of available resources to serve children who are in need of developmental interventions. Although the framework for early intervention service exists, we agree that additional research would help optimize the ultimate impact of our developmental and autism-specific screening recommendations.

Vidya B. Gupta, MD
On behalf of the AAP Autism Expert Panel
Member, AAP Autism Expert Panel
Department of Pediatrics
New York Medical College
Valhalla, NY 10029

Paul H. Lipkin, MD
On behalf of the AAP Policy Revision Committee on Developmental Surveillance and Screening
Chair, AAP Policy Revision Committee on Developmental Surveillance and Screening and AAP Council on Children With Disabilities
Division of Neurology and Developmental Medicine
Kennedy Krieger Institute
Baltimore, MD 21205
Department of Pediatrics,
Johns Hopkins University School of Medicine
Baltimore, MD 21287

REFERENCES

doi:10.1542/peds.2007-0772
Marketing Foods and Beverages: Why Licensed Commercial Characters Should Not Be Used to Sell Healthy Products to Children

To the Editor.—

We applaud the American Academy of Pediatrics on the policy statement *Children, Adolescents, and Advertising*. Current advertising practices certainly deserve the careful scrutiny of all who care about our nation’s children, as well as our nation’s future. The American Academy of Pediatrics statement presents a critical opportunity to begin to change current advertising practices.

An area that the policy statement did not address but is of equal concern is the use of licensed commercial characters to sell healthy foods and beverages to children. A 2005 Institute of Medicine report entitled *Food Marketing to Children and Youth: Threat or Opportunity?* recommended “that licensed commercial characters [be] used only to promote foods and beverages that support healthful diets for children and youth.” In addition, in an effort to build consensus on how to combat childhood obesity, the Federal Communications Commission recently formed a task force entitled *Media and Childhood Obesity: Today and Tomorrow*. At the California Childhood Obesity Conference in Anaheim, California, this past January, representatives from a number of the organizations participating on the Federal Communications Commission task force presented their perspectives on the ways in which media contribute to and may help solve the obesity epidemic. One theme that was echoed by all of the different voices and was particularly alarming was support for the Institute of Medicine’s recommendation that licensed commercial characters should be used to sell healthy choices to children.

We know that children connect with particular characters and will follow those characters out of loyalty to them. However, it is important to carefully consider the potential ramifications of using characters to manipulate children to desire any particular product. Persuasion of children, whether for unhealthy or healthy ends, undermines the authority of parents to make decisions on their children’s behalf and exploits children’s healthy development and their positive, natural tendencies toward loyalty and belonging.

The following are some specific reasons that licensed commercial characters should not be used to sell any food or beverage products to children.

1. There is evidence that children cannot distinguish between advertising messages and program content before the age of 8. The exploitation of this developmental stage cannot be justified even for the noblest of causes.

2. Teaching children to be media literate is critical to their healthy growth and development. Using characters to encourage them to use specific health-related products sends conflicting messages.

3. When commercial characters are placed on food packaging, we are not only using the character to sell the food, but we are using the food to sell the character. Refrigerators and pantries stocked with food stuffs, healthy or unhealthy, do not need to become billboards for television shows or movies.

4. Parents trying to ensure their children’s health should not have to make concessions to commercialization regardless of how healthy a food is deemed to be by others.

5. Many commercial characters do not serve as universally positive role models. Using characters with negative traits to sell healthy behaviors may fix one problem but reinforce others.

6. Some marketing has been designed to encourage children to nag their parents and “act out” if they do not get what they want. This is not a behavior we want to reinforce, even if the product that children desire is considered to be healthy.

7. Finally, the decision regarding whether a particular product is healthy or unhealthy is extremely complicated. Condoning the use of characters to sell “healthy” products opens a Pandora’s box in defining terms. The food industry remains a step ahead of health professionals in designing unhealthy food and beverage products to meet “health” guidelines. We must protect children from corporate interests that aim to define health and determine dietary behavior.

Educators, advocates, policy makers, corporations, and the media are working together to improve children’s physical activity and dietary behaviors. These critical actions will impact the future of our nation’s children. We should support children’s programs that promote healthy eating and physical activity during the programming itself, including those in which characters model these behaviors. These messages should be general and universally accepted and should not promote particular products.

A campaign to educate young children about healthy choices and even to market healthy foods and beverages to parents of young children should be undertaken by health experts with the financial support of the government, foundations, and responsible corporations. This campaign should not include the use of licensed commercial characters to promote any particular food or beverage items.

Although it will be difficult to change marketing and advertising practices in our country, health professionals and advocates who work on behalf of children must be
wary of the use of commercial characters to sell even healthy products to children.

By respecting our children, freeing them from commercial interests, and allowing them to grow and develop healthfully, we can reverse the current obesity epidemic and enhance the opportunity for the healthy, vibrant, and productive future that all children deserve.

Wendi Gosliner, MPH, RD
Center for Weight and Health
University of California
Berkeley, CA 94720

Kristine A. Madsen, MD, MPH
Department of Pediatrics
University of California
San Francisco, CA 94118

REFERENCES

doi:10.1542/peds.2007-0724

In Reply.—

As lead author of the recent American Academy of Pediatrics policy statement “Children, Adolescents, and Advertising,” I very much appreciate the authors’ endorsement and concerns. Obviously, in an ideal world, we would not advertise anything, good or bad, to children. As the authors (and our statement) noted, children under the age of 8 cannot distinguish between what is advertising and what is program content. This was the finding of the Federal Trade Commission 25 years ago when they seriously considered banning all advertising aimed at young children at that time. Unfortunately, they could not figure out a practical way to do that. Other countries (Australia, Netherlands, Sweden), however, have figured it out, and they have lower rates of childhood obesity. In Britain, a ban on advertising for food or beverages that are high in fat, salt, or sugar in programming for children and teens is currently being considered. Any program that attracts an average proportion of viewers under 16 higher than in the general population would be affected, which, of course, is the way to solve the regulation problem that the Federal Trade Commission wrestled with 25 years ago.

Unfortunately, in this country, it seems that when public health comes into conflict with capitalism, the latter always wins out. Given that we have the best Congress money can buy, it seems unlikely that any significant legislation will appear anytime soon, especially because the food industry has nearly 100 different lobbying groups in Washington, DC. Only the American Academy of Pediatrics, the American Psychological Association, and a few other groups seem to be lobbying on behalf of America’s children and teenagers. We are outspent and out-gunned. As such, let’s agree to work on bad advertising first and worry about the “good” advertising later. However, I agree that we should not be advertising to young children at all with either good or bad advertising.

Víctor C. Strasburger, MD
Department of Pediatrics
University of New Mexico School of Medicine
Albuquerque, NM 87122

REFERENCES

doi:10.1542/peds.2007-0825

Mycoplasma pneumoniae Encephalitis and Reactivation of Herpes Simplex Virus

To the Editor.—

Elbers et al. recently reported on 16 patients with encephalitis and evidence of herpes simplex virus (HSV) infection. Four patients also had evidence of an acute infection with Mycoplasma pneumoniae (polymerase chain reaction– or immunoglobulin M–positive cerebrospinal fluid [CSF]). In 3 of those cases (cases 7–9), it is more
likely that *M pneumoniae* was the primary pathogen causing the encephalitis. *M pneumoniae* is a common pathogen in children with encephalitis and accounts for up to 10% of cases in prospective studies. Similar to HSV, it causes local high signal density lesions on T2-weighted MRI and electroencephalographic changes including slow-wave activity. *M pneumoniae* is not a contaminant in the CSF and, unlike HSV1 or HSV2, would not emerge from reactivation. Case patients 6 and 7 had no rise in anti-HSV complement-fixation-test titers, and case patient 9 remained immunoglobulin G–negative, which is against a recent systemic HSV infection and more in favor of a local reactivation. A local reactivation may also be the reason why the polymerase chain reaction test for HSV was initially negative in case 9. *M pneumoniae* elicits high levels of interleukin 6 in the CSF during encephalitis, which has been shown to reactivate HSV. Treatments of *M pneumoniae* encephalitis has been controversial, because antibiotics used like macrolides may not achieve adequate CSF levels. Chloramphenicol may be a better choice, but there have been no randomized, controlled trials that used this drug for that purpose. Experience with steroid treatment against an autoimmune component of mycoplasma encephalitis is limited. Future experimental studies need to clarify whether it is the reactivation of HSV or the primary mycoplasma infection that causes the encephalitis and neuronal damage in these cases.

Michael Eisenhut, MD
Department of Pediatrics
Luton and Dunstable Hospital
NHS Foundation Trust
Luton LU4 0DZ, United Kingdom

REFERENCES


doi:10.1542/peds.2007-0616

In Reply.—

We thank Dr Eisenhut for his interest in our prospective 12-year study of herpes simplex encephalitis in children. Because of the standardized comprehensive prospective microbiologic investigations performed on all children who were admitted to our institution with acute encephalitis, evidence of infection with >1 potential pathogen is often observed. This approach is very different from retrospective reviews of cases in which the final diagnosis is the point of entry into the study. In a recent review of our data, approximately one fourth of the cases showed evidence of infection (of variable strength) with ≥2 pathogens. In situations such as this, it is often not possible to determine with certainty which pathogen is the causative agent. Furthermore, in some cases, >1 of the potential pathogens may be playing a role.

In regard to the specific cases mentioned by Dr Eisenhut, we disagree with his conclusion that *Mycoplasma pneumoniae* is the more likely pathogen. We believe that most experts would consider the detection of herpes simplex virus (HSV) in the cerebrospinal fluid (CSF) by polymerase chain reaction (PCR) as strong evidence of causality irrespective of HSV serologic results or microbiologic evidence of infection with other potential pathogens. The failure to detect a rise in complement fixation titer to HSV (cases 7 and 8) between acute and convalescent sera could be a result of the well-known limited sensitivity of this assay. In regard to case 9, it is possible that the convalescent sample was obtained too early (51 days) for HSV seroconversion to have occurred.

In 1 prospective study of HSV seroconversion, the median time between onset of genital tract symptoms and seroconversion was 56 days (interquartile range: 16–121 days). In addition, there is a paucity of data on the sensitivity of HSV serology in the setting of HSV encephalitis. We acknowledge that when both HSV and *M pneumoniae* are detected by PCR in the CSF, it is difficult to assign causality solely to 1 pathogen. To avoid subjective assignment of causality, we set rigid inclusion criteria for herpes simplex encephalitis; these criteria are detailed in the methods section of our article.

Dr Eisenhut refers to “local reactivation” of HSV as a possible explanation for the delayed detection of HSV in the CSF of case patient 9. We do recognize that local reactivation (not causally related to the encephalitis) may be an explanation for some patients with positive HSV CSF PCR results but believe this to be unlikely in this case because of the aforementioned stringent criteria used to define HSV encephalitis. We did exclude several cases in which HSV was detected in the CSF by PCR because of a clear alternative diagnosis or absence of typical CSF, electroencephalographic, or neuroimaging findings, again in accordance with preselected exclusion criteria. As discussed in our article, we support a full course of therapy for possible HSV encephalitis for children in whom HSV is detected in the CSF by PCR irrespective of the suspected cause of the encephalitic process. In addition, it has been shown that the detection of HSV in the CSF by PCR of patients with HSV encephalitis can be delayed. Hence, HSV should not be excluded as a possible cause of encephalitis solely on the basis of a
single negative CSF PCR assay, particularly if the lumbar puncture was performed during the first 48 to 72 hours of illness.

We agree that *M pneumoniae* is an important cause of acute encephalitis in children and that additional research into the role of this pathogen in neurologic disease is needed. In a previous publication pertaining to *M pneumoniae* and acute childhood encephalitis, we stratified cases into probable, possible, and indeterminate categories in accordance with the strength of microbiologic evidence implicating *M pneumoniae* as the cause. Accordingly, *M pneumoniae* was considered to be the probable cause of 6.9% of the acute childhood encephalitis cases. As indicated in that article, cases in which the only evidence of *M pneumoniae* infection was serologic were often associated with more compelling evidence of infection with other potential pathogens. This observation suggests that many of the positive immunoglobulin M assays for *M pneumoniae* were probably falsely positive, perhaps as a result of cross-reactivity of *M pneumoniae* antigens and human brain antigens or other inherent limitations in the specificity of serologic tests for *M pneumoniae*.

Determining the etiology and pathogenesis of acute childhood encephalitis remains a major hurdle to the appropriate management of this condition. Ultimately, new innovative diagnostic and treatment strategies will be needed to improve the outcome of this often devastating entity.

Ari Bitnun, MD, MSc
Division of Infectious Diseases
Department of Paediatrics
Jorina M. Elbers, MD
Daune MacGregor, MD
Division of Neurology
Department of Paediatrics
Raymond Tellier, MD, MSc
Susan E. Richardson, MD
Department of Paediatric Laboratory Medicine
Hospital for Sick Children and University of Toronto
Toronto, Ontario, Canada MSG 1X8

REFERENCES


Give Synagis Via Home Care: A Reply

To the Editor,—

I would like to expand on the issues raised by Dr Gold-enring in his January 2007 letter to the editor.1 Dr Gold-enring suggests that US health care authorities should consider the administration of Synagis (palivizumab) via home care.

In my home country of the Netherlands, and similar to clinical practice in the United States, at-risk candidates for respiratory syncytial virus (RSV) prevention with Synagis include those with prematurity, chronic lung disease, and congenital heart disease. When their at-risk infants are healthy enough to be discharged from the hospital, Dutch parents receive comprehensive instruction on proper hygienic measures to minimize infant exposure to common viruses such as RSV. In many cases in the Netherlands, at-risk infants receive their monthly Synagis injections from qualified nurses in the safety of their own homes. Specifically, Synagis home care in the Netherlands is administered by an independent third party, Klinerva.

Home administration of Synagis is highly consistent with the education we give to parents on protecting their at-risk infants from viral attack. Why should at-risk infants be forced outdoors during the cold winter months, exposed to the many secretions of sick children in communal waiting areas within hospitals, solely for the purpose of receiving their monthly injection of Synagis?

Interestingly, a multicenter (22-hospital) study was recently conducted in the Netherlands to assess the via-
ability and effectiveness of home-care administration of Synagis. A total of 334 infants (40% male, 60% female) who received Synagis via home care with Klinerva during the 2005/2006 RSV season were followed. As part of the study, parents were asked to rate, on a scale from 0 to 10, their overall experience with the home-care program. Not surprisingly, the program scored well with an average score of 8.8 of 10.

In closing, I respectfully suggest that we, as clinicians, need not cling to unnecessary habits of the past, especially when doing so has the potential to expose a vulnerable infant to needless additional risk. And, let us not forget the parents. Five fewer visits to the neonatologist or pediatrician during the RSV season must certainly be a welcome relief given parents’ busy schedules. In today’s world, home-care administration of Synagis should be embraced.

Johannes A. Schipper, MD, PhD
St Antonius Hospital
3430 EM, Nieuwegein, Netherlands

REFERENCE

doi:10.1542/peds.2007-0715

In Reply.—
I thank Dr Schipper for his supportive comments. I am not surprised to find that the Netherlands, a country with a well-designed national health system, would encourage administration of Synagis in children’s homes.

Again, I call on the American Academy of Pediatrics Committee on Infectious Diseases to consider this issue and change its guideline to specify home administration of Synagis during respiratory syncytial virus season.

I also believe it is time that we consider administering all immunizations to these most vulnerable newborns at the same home visits. These visits will also offer, if we wish, an opportunity to assess the home environment of these maximally stressed families with very small, and sometimes very ill, infants, with a view to preventing child abuse, among other possible positive spin-offs.1

It has been postulated that the automobile was a huge factor in changing how we provide medical care in the United States.2 People could visit us easily. Hence, office-based practices developed. However, now we must return full-circle to realize that we can go back out to homes, if not by ourselves then at least through the services of our highly trained nursing colleagues.

As a medical director, I have participated in ongoing trials of physician- and nurse practitioner–based home care for the most vulnerable elderly patients. None of the busy physicians who cared for these very ill patients raised any objection to the care rendered as long as communication was good. Therefore, like Dr Schipper, I see no reason why busy pediatricians would feel differently. I suggest that home nursing programs for vulnerable newborns and infants do not diminish or interfere with the concept of the medical home; they merely extend it.

John Goldenring, MD, MPH, JD, FAAP, FSAM, FCLM
San Diego, CA 92131

REFERENCES

PULLING THE PLUG ON ENTERTAINMENT INDUSTRY RATINGS

To the Editor.—
As recently as 2006, US Congressional hearings have explored problems with the entertainment industry’s rating of violent content in video games, movies, and television (TV).1,2 The ratings developed by the media industry classify glamorized media violence as appropriate for young children—a position that is inconsistent with the scientific consensus about the negative effects of media violence and contradicts American Academy of Pediatrics (AAP) media policy. I suggest that the AAP discontinue its endorsement of parents’ use of the entertainment industry ratings.

As the established leader among medical organizations in children’s media issues, the AAP plays an important role in helping parents make effective decisions about their children’s media use. The AAP’s advice to parents regarding violent media content is especially important, because the pathologic effects of such media have been established within the scientific community.3

Problems related to children’s exposure to media violence, including increased real-life violence, were articulated by the AAP and other health organizations in the “Joint Statement on the Impact of Entertainment Violence on Children.”4 This statement identified particular risks associated with media violence that is glamorized. Media violence is glamorized by depicting it in a positive light, such as when it is rewarded or perpetrated by children’s role models.

Entertainment industry ratings do not accurately convey the risks associated with children’s exposure to glamorized media violence. A study of video games that the industry has rated E (may be appropriate for children aged 6 years and up) found that 60% of the games tested required players to hurt characters in return for rewards or advancement.3 TV and movie ratings similarly indi-
cate that media with high levels of glamorized violence are suitable for young children.\(^5\)\(^7\)

Problems with entertainment industry ratings have led the AAP to recommend, in its “Media Violence” policy statement, that a new rating system be developed that helps parents make healthy media choices for their children.\(^1\) Similarly, organizations including the National Institute on Media and the Family, the American Medical Association, and the American Psychological Association have said that the current ratings are not protecting children’s health, and they advocate a new ratings system.\(^8\)

The entertainment industry has resisted changing its rating system, most likely because of the advantages these ratings provide the industry. Because ratings may have an impact on the revenues gained from a particular media product, it is in the industry’s interests to apply its own ratings.\(^9\) The current rating system has also served the industry as a public relations tool when defending itself for producing violent products\(^10\) and marketing them to children.\(^11\)\(^12\)

Despite its own policy and the significant problems with the current rating system, the AAP suggests (in a public education brochure for parents) that parents use entertainment industry ratings as an effective means for guiding their children’s media use. The AAP instructs, “Look for ratings and warning labels. Use them to make smart decisions about what your child sees and hears.”\(^13\) Moreover, the AAP implicitly endorses the entertainment industry ratings by placing the video game, movie, and TV rating icons and descriptions on its Web site for use by parents.\(^13\)\(^16\) For example, on the AAP’s “Entertainment Rating System: Video and Computer Games,”\(^15\) the icon for the video game industry’s E rating is replicated followed by the statement, “Suitable for ages 6 and older.” This rating suggestion contradicts the AAP’s counsel in its “Media Violence” policy statement, which recommends “avoiding violent video games in homes where they may be observed or played by young children.”\(^3\)(p1224)

The fundamental problems with the entertainment industry’s rating of media violence suggest that the AAP withdraws its recommendation that parents use the rating system. The AAP could help parents by removing industry rating icons and descriptions from the AAP educational materials for parents and its Web site; informing them that the current ratings are not evidence based and do not account for the destructive impact of glamorized media violence; and directing them toward evidence-based resources that effectively describe the hazards of media violence. AAP Web sites and brochures could also outline the potential harm associated with children’s viewing media violence that is rewarded, committed by children’s role models, or otherwise glamorized.

AAP communications with parents about the failings of entertainment industry ratings could help pave the way for the development of valid, evidence-based media ratings founded on health outcomes, which pediatricians could confidently recommend to parents. When the public understands that the current ratings do not accurately describe the health risks of media use, parents and other advocates for children can join with the AAP as a positive force for change.

Richard Freed, PhD
Department of Mental Health
Kaiser Permanente Medical Offices
Antioch, CA 94509

REFERENCES
This challenge to American Academy of Pediatrics recommendations about using media ratings is completely consistent with best practice of pediatrics. Pediatricians, who translate complex science into strategies for optimizing child health, have led public concern over the effects of media on children for more than 2 decades. What does research show about media violence? The Center on Media and Child Health has catalogued 956 scientific articles (www.cmch.tv) that provide nearly unanimous evidence that exposure to media violence contributes to elevated fear and anxiety, sleep disturbances, desensitization to human suffering, and increases in aggressive thoughts and behaviors.1

Media ratings are neither based in science nor assigned by child development experts. Applied by entertainment industry employees to inform consumers about “inappropriate” content, ratings estimate publicacceptability, not potential to influence health and safety. Media producers know their age-based ratings are effective tools of marketing and self-protection. Children will seek out content aimed at more mature audiences. Ratings of any quality allow producers to dissuade legislators from restrictions and point to parents as those responsible for choosing what their children view. Allowing producers to rate their own product is akin to allowing butchers to certify the safety of meat. In no other aspect of practice would pediatricians recommend abandoning science and relying on the advice of those who stand to profit from children’s use of a health-affecting product.

Physicians demand rigor, reliability, and validity of the science that guides medical decision-making. We want information that is honest, complete, and applicable to issues of critical concern. Industry ratings meet none of these standards. Compared with parent ratings of film and television, industry ratings varied by as much as 50%, all in the direction of permitting younger viewers.2 On average, 1 decade of “ratings creep” permitted R-rated content to be rated PG-13.3 Every G-rated animated feature made between 1937 and 1999 portrays violence.4 Video game ratings ignore the most severe content at all rating levels.5,6 Beyond violence, research has shown smoking in movies to be a powerful independent influence on youth initiating and maintaining tobacco use,7 yet smoking portrayals have no influence on ratings.8

Children are exposed to and powerfully affected by media. Pediatricians should encourage informed decision-making on safe media use. When this recommendation was first made, only industry ratings were available. Today there are other sources of media content information, several of which are more accurate, and certainly less self-serving, than the industry ratings. As a society, we have demanded inspection and certification of what we feed our children’s bodies, but we have no credible system for ensuring the safety of what we feed their minds. As pediatricians, we need something better than industry ratings with which to guide parents and children toward optimizing their physical, mental, and social health.

Michael Rich, MD, MPH
Center on Media and Child Health
Children’s Hospital Boston
Harvard Medical School
Harvard School of Public Health
Boston, MA 02115

REFERENCES
1. American Academy of Pediatrics, Committee on Public Educat-
2. Walsh DA, Gentile DA. A validity test of movie, television, and 
3. Thompson KM, Yokota F. Violence, sex, and profanity in fil-
2004;6:63
4. Yokota F, Thompson KM. Violence in G-rated animated films 
[published correction appears in JAMA. 2000;284:567]. JAMA. 
2000;283:2716–2710
JAMA. 2001;286:591–598
6. Haninger K, Thompson KM. Content and ratings of teen-rated 
video games. JAMA. 2004;291:856–865
7. Sargent JD, Beach ML, Adachi-Mejia AM, et al. Exposure to 
movie smoking: its relation to smoking initiation among US 
8. Charlesworth A, Glantz SA. Smoking in the movies increases 
1528

Adverse Events in the Manipulation of Pediatric Patients: Flaws in a Systematic Review

To the Editor.—

There is no question that the issue of safety in spinal manipulation, as with any medical intervention, needs
to be thoroughly addressed if it is to be considered an effective and broadly used modality of health care. That said, numerous articles, often containing numerous methodologic flaws,1–3 together with responses4–6 have appeared in the literature addressing the issue of manipulation (usually cervical) in adults. The recent article by Vohra et al7 was one of the first to attempt to more comprehensively cull the observational studies as well as the randomized clinical trials to date in assessing the safety matter of manipulation as it pertains to children. It correctly stated that there is useful information to be gleaned from the case studies as well as the more rigorous clinical trials.

That said, however, there are flaws significant enough to invalidate 2 of the authors’ key objectives:

### COMPLETENESS AND ACCURACY OF THE SYSTEMATIC REVIEW

The authors’ claim that they thoroughly searched the literature is threatened in their introduction, where they completely omitted any reference to the clear benefits of spinal manipulation in treating otitis media, as reported in 2 large cohort studies published in the indexed literature by Froehle8 and Fallon and Edelman.9 Of far more gravity, however, is the fact that they reported a series of 14 cases of direct adverse events and that “each case involved a chiropractor and was reported in the United States.” From 1 of the references cited, however, the treating doctor was most decidedly not a chiropractor but, rather, a physical therapist,10 most likely practicing not in the United States but in Germany, where chiropractors are outnumbered by >150-fold by medical doctors legally performing >14 million treatments yearly in the spinal area, practicing what they call “chiropractic treatment.”11

Another citation of a “direct adverse event” involved a feasibility study for enrolling children affected by otitis media into a clinical trial involving spinal manipulation by chiropractors. The presumed seriousness of this “adversity” was reported as follows:

> There were no serious side effects as a result of either the active or placebo chiropractic treatments. One parent in the active treatment group reported that their child had some mid-back soreness after one treatment that had resolved after a few days, and another child was reported by the parent as being irritable for a short time after treatment. One parent of a child in the placebo chiropractic group reported excessive crying by the child after treatment.12

Besides seeming to be a truly minor event that is almost misleading to classify as “adverse,” the difference of these transient setbacks between active and placebo chiropractic treatment groups seems sufficiently narrow as to call into question data such as these. Simply speaking, are these presumably negative effects even valid enough to have been admitted into the Vohra et al study?

The most troubling reference in this study pertains to 3 cases of delayed treatment for the life-threatening conditions of meningitis and embryonal rhabdomyosarcoma in 1 of the articles included in the review. Besides not even being a direct consequence of spinal manipulation per se, the delayed diagnoses referred to by the authors are not even given any mention in 1 of the citations presented. Instead, this particular article focused solely on cerebrovascular incidents that follow chiropractic treatment.13 Oddly, the authors seemed to retract their own implication by stating as 1 of their study limitations that “our search strategy did not allow for systematic identification of indirect adverse events (eg, delayed or missed diagnoses).” If that was, in fact, the case, why was it discussed earlier in their article as indicated above?

These examples are sufficient to call into question whether the Vohra et al study even accomplished its minimal objective, which was to present an accurate and balanced assessment of the literature pertaining to the adverse events associated with spinal manipulation, presumably by chiropractors as set out by the authors in their introduction.

### THE NEED TO QUANTIFY THE RISK ASSOCIATED WITH MANIPULATION

After discussing the supporting and dissenting viewpoints on spinal manipulation in children, the authors rightfully stated that “[t]here is an urgent need to quantify the risk associated with spinal manipulation in children.” In addition to including events that could hardly be considered “adverse” by most accounts, the authors concluded that the reports sampled in their study did “not provide information on the incidence of adverse events because of the lack of data regarding the total number of manipulations provided (ie, denominator data).” Thus, the primary purpose of the Vohra et al study seems to remain unanswered. In addition, there are some lesser concerns.

**Categories of Classification of Adverse Events**

The authors’ categorization of the adverse events into 4 categories (adverse, severe, moderate, and minor) has not been validated. Furthermore, categorizing hospitalization as severe per se seems to ignore the fact that nonhospitalized occurrences could be more intense and disabling at times. Finally, the authors categorically assigned a fifth category (delayed diagnosis or treatment) as a “moderate to severe” adverse event. However, there seems to be no systematic or rational basis for this last classification.
Unreported Adverse Events

The authors rightfully reported that underreporting of adverse events in branches of medicine outside of chiropractic may be widespread, and cited a study that suggested that <10% of serious adverse events are reported as drug reactions. In reporting major incidents in hospitals, 1 Harvard School of Public Health professor indicated that many medical errors are barely investigated because of the lack of resources. In fact, only 20 states require the reporting of medical errors at all. Spinal manipulation, by comparison, would seem to involve far fewer numbers with overall less severity than what has occurred in the medical arena. Such is not to exonerate chiropractic from responsibility in safety matters but simply to make it clear that the authors have significantly weakened their case for this study, having made chiropractic safety a more pressing matter by stating that “[g]iven the large numbers of children who have received spinal manipulation during the decades assessed by our search strategy, adverse events resulting from spinal manipulation are either remarkably rare [italics mine] or underreported.”

Specification of Maneuvers

“Rapid and/or strong rotational maneuvers” were specified by the authors as having precipitated 2 of the 4 serious adverse events. Yet, 1 of the studies used a physical therapist and not a chiropractor as erroneously stated by the authors, as pointed out earlier. The other study they cited is dated almost 30 years ago, which leads one to suspect, or at least hope, that significant enlightenment and progress has been made since then.

Spontaneity of Events

By citing a study that suggests that >1 in 3 spontaneous case reports of rare or uncommon adverse events is unlikely to be coincidental, the authors create the impression that the same holds true for spinal manipulations. Recently, by far the majority of serious adverse events associated with spinal manipulation have been linked to the chiropractic treatment of the cervical region leading to vertebral artery dissections. However, it has been pointed out that the rate of spontaneous arterial dissections ranges from 1 to 3 per 100,000, which may be as much as 10 times the rates of arterial dissections that have been attributed (and not even linked causally) to spinal manipulation. Therefore, the Vohra et al study may have created a misleading impression that spinal manipulation is a more serious problem than has been currently regarded. In reality, the responses to the topic of cervical manipulation of the mass media in particular have sometimes approached hysterical dimensions in reaction to the flawed investigations cited involving adult patients.

CONCLUDING REMARKS

From several vantage points, the Vohra et al study fell short of the mark of having achieved its objectives or even maintaining its credibility as a systematic review of the literature pertaining to cervical manipulation. For these reasons and its inexplicable inaccuracies within a framework of peer review that is designed to eliminate such fallacies, the Vohra et al study can be greeted only with the most extreme skepticism.

Anthony Rosner, PhD, LLD
Foundation for Chiropractic Education and Research
Brookline, MA 02446

REFERENCES

The issue of chiropractic pediatrics is of much greater concern than the specific dangers outlined by Vohra et al. Pediatric chiropractic is taught in every chiropractic school in North America. The main textbook, *Pediatric Chiropractic*, is 775 pages long; it lists every pediatric condition, from ear infections to epilepsy to attention-deficit disorders, that is diagnosed by chiropractors, and a chiropractic treatment, mostly highest neck manipulation, is proposed for each. Chiropractic authorities offer courses all over North America whereby a chiropractor can be certified as a specialist in pediatrics. The title granted is the FICPA, which stands for Fellow of the International Chiropractic Pediatric Association (see www.icpa4kids.com, www.icapediatrics.com, and www.chiropediatrics.com).

Chiropractic pediatrics is fully supported by every state and provincial chiropractic licensing board and by the National Board of Chiropractic Examiners. All of them consider pediatric chiropractics to be acceptable as a standard of care and support pediatric chiropractic treatment fully. Our surveys in Canada, which use government data, show pediatric chiropractic to be a $50 million/year enterprise, which would be $500 million annually in the United States.

Chiropractic is defined by a belief in vertebral subluxations. The most important subluxations are those of the joints between the base of the skull, the atlas and the axis. No matter what the patient complains about, the highest neck of the infant or child is almost always manipulated.

This subluxation belief claims that a chiropractor, by use of his or her hands alone, can diagnose vertebral bones as being out of place and determine that this displacement is causing pathology via the autonomic nervous system to various organs of the body, and this can all be corrected by manipulating the subluxated vertebrae.

Considering that the nuchal area of an infant is 2 to 3 inches and is covered by fat, it is clear that pediatric chiropractic, and indeed all of chiropractic, is an anatomic and physiologic impossibility that has never been true and can never be true.

Chiropractors learn to play the research game (see www.icpa4kids.com, www.icapediatrics.com, and www.chiropediatrics.com). They have published studies on ear infections that have been cured by chiropractors who never looked into the ear or colic that was cured by the sacral bone being manipulated. Although scientific research is based on finding treatments to cure disease, chiropractic research goes backward; it is a treatment in search of a disease. The same treatment cures everything, so it must work.

Chiropractic is dangerous for infants and children because, as was stated by the Chief of Pediatrics of Canadian Hospitals, it is ineffective and useless. The key-stone of the practice of pediatrics is to provide scientific information to parents so that they can properly care for their children. Chiropractic does the opposite. Pediatric chiropractic lecturers teach chiropractors how to shift credibility from the pediatrician to the chiropractor.

The fallout of this exists; we just often do not know about it. In addition, chiropractic involves excessive and totally unwarranted use of radiation and is antimunization. As stated by the Chiefs of Pediatrics of Canadian Hospitals, “Chiropractic use of x-rays of infants and children to diagnose so called vertebral subluxations is un-scientific and of no value whatsoever. Without any benefit to the child, these x-rays can contribute to the risk of cancers and genetic damage. Parents should never allow their children’s spines to be x-rayed by chiropractors.”

There are specific actions that should be undertaken. The American pediatric community should follow the lead of the Canadian pediatricians and neurologists who have helped stop almost all government funding of chiropractic in Canada. All government financial support for any chiropractic school that teaches subluxation theory for those in the pediatric age group (birth to 18 years of age) must be stopped. Although chiropractic schools have failed to gain university recognition and are privately funded, the graduates do enjoy use of the title “Dr” and receive insurance and other payments. Public health authorities can demand that for chiropractors to continue to benefit in any way, pediatric chiropractic subluxation theory must not be taught. Most importantly, all the chiropractic licensing bodies that endorse pediatric chiropractic subluxations as an acceptable standard of care should be investigated by public health authorities and called on to declare pediatric chiropractic as being not acceptable.

The fact remains that 95% of chiropractic treatments of infants and children involve nothing at all being done except gently turning the head sideways by pushing on the back, even through the clothing. Although harm has come to children from spinal manipulation (see the Neck911USA database at www.neck911usa.com), this is not where the focus of concern should be. The issue is much wider and involves direct patient education from pediatricians’ offices on this issue as well as educational and legislative reform as suggested above.

Murray S. Katz, MDCM
Tiny Tots Medical Centre
Dollard des Ormeaux, Quebec, Canada H9B 2Y1
To the Editor.—

The American Chiropractic Association (ACA) commends Vohra et al1 for their efforts to learn more about the types and frequencies of adverse reactions associated with pediatric spinal manipulation. We are concerned, however, that the recent study published in Pediatrics does little to add to the scientific literature and will only create undue alarm among parents.

Although Vohra et al found a limited number of reports that identified adverse reactions among pediatric chiropractic patients over the past 60 years, an objective reader could question how these statistics compare to the number of reactions among the same age group that is associated with traditional medical care. Nowhere does the study indicate that analgesics, antibiotics, or any other treatment is safer than spinal manipulation for neuromusculoskeletal conditions. In fact, recent reports suggest that prescription and over-the-counter medications are inherently riskier than conservative spinal manipulation.2–7

Chiropractic care is always adapted to the individual patient. It is a highly skilled treatment and, in the case of children, very gentle. Even so, doctors of chiropractic routinely inform patients to expect some soreness after their first adjustment. The current literature documents that some patients may experience minor discomfort after spinal manipulation that usually fades within 24 hours.8 The chance that chiropractic manipulation will result in a serious reaction in adult patients is remote, ranging from 1 in 1 million to 1 in 5.85 million manipulations.9,10 For the majority of patients looking for drug-free alternatives, chiropractic manipulation has been shown to be safe and effective for neuromusculoskeletal pain.

Furthermore, the authors’ recommendation for parents to initially consult with their child’s medical physician before seeking the professional services of a doctor of chiropractic is unwarranted. Chiropractors are uniquely trained and licensed in every state as primary care providers. They are aware of the types of conditions that will respond to their care, and they can also recognize those conditions that require referral to other health care providers/specialists.

Evidence from many types of experimental studies (basic and clinical, comprised of randomized, controlled trials and cohort and case studies) provides a promising basis with which to consider chiropractic management for such childhood conditions as otitis media, asthma, and enuresis. The current data suggest that further exploration of alternative treatments, such as spinal manipulation, would benefit patients and parents alike.11

It is the ACA’s long-standing belief that doctors of chiropractic, in concert with other health care providers, can play an important role in the health of children. The care and treatment of children is included in the chiropractic profession’s scope of practice and taught in and through accredited colleges of chiropractic. These colleges teach physical examination, diagnostic procedures, and patient management skills necessary for the quality care of children.

Although chiropractic care has been proven to be beneficial to patients of every age, the ACA believes that doctors of chiropractic are duty-bound to diagnose health conditions brought to their attention and treat the patient if the condition is within their stated scope of practice; otherwise, they should refer the patient for more specialized care.

The ACA encourages additional studies that further evaluate pediatric spinal manipulation and the possibility of a rare adverse event. Unfortunately, the Vohra et al study, although well intended, leaves concerned parents and pediatricians with more questions than answers.

Richard Brassard, DC
President, American Chiropractic Association
Arlington, VA 22209
Texas Chiropractic College
Pasadena, TX 77505

REFERENCES

doi:10.1542/peds.2007-0299

PEDIATRICS Volume 119, Number 6, June 2007 1265
corrections. We would like to address some of the other concerns raised.

The objectives and methods of this review seem to have been misinterpreted by some. The purpose of our systematic review was to assess the safety of pediatric spinal manipulation, not its efficacy or the safety of other interventions. These topics were appropriately omitted from our article. In addition, our review was not limited to cervical manipulation, as Rosner seemed to indicate. Because we could not locate validated adverse-event categories, we chose categories/definitions that are consistent with those used by the National Institutes of Health. By definition, minor events had to be self-re-solving and not require additional medical care. Each of the minor adverse events reported in our article met the a priori definition. We did not think it was reasonable to omit the delayed diagnosis/treatment that was revealed through our systematic review, and we, therefore, applied the same categories to classify the resulting harms.

We actively collaborate with a variety of complementary and alternative medicine institutions, including a national chiropractic college. Our recent research has identified a need for more pediatric training for chiropractors, an opinion voiced by chiropractors themselves. In fact, 97% of 180 responding chiropractors recommended increased pediatric training. Our systematic review corroborates that children with serious underlying disease may suffer delayed diagnosis or treatment. It seems only prudent to encourage parents to discuss their child’s health and health care practices with their child’s physician.

Because of the paucity of data, risk estimates or causal inferences were not possible. Serious adverse events are often rare and require population-based exposure before they can be detected. Previous experience has shown that multiple spontaneous reports of a given adverse event are unlikely to be a result of chance alone. It is not possible to identify the incidence rate of adverse events without better-quality data. Lack of reported adverse events about pediatric spinal manipulation should not be interpreted as a confirmation of safety, nor can it be deemed prima facie evidence of underreporting. Our systematic review clearly illustrates the gap in knowledge on this topic. We urge multidisciplinary collaboration to prospectively quantify risks associated with pediatric spinal manipulation. A multidisciplinary approach by all those who provide spinal manipulation (eg, chiropractors, physicians, physical therapists) should be encouraged to minimize bias and maximize “buy-in” from the disciplines involved. Health care providers cannot seek informed consent, nor can parents provide it, if the risks are unknown. Our intent was to highlight the need for additional research by contrasting it with the scant available data on this topic. If we are to learn from conventional medicine, that lesson should be that patient safety is not assured by good intentions alone.

Sunita Vohra, MD, MSc, FRCP C
Bradley C. Johnston, ND
Complementary and Alternative Research and Education Program
Department of Pediatrics
University of Alberta
Edmonton, Alberta, Canada T6G 2E1

Kim Humphreys, DC, PhD
Graduate Education and Research
Canadian Memorial Chiropractic College
Toronto, Ontario, Canada M2H 3H1

REFERENCES

doi:10.1542/peds.2007-0503

In Reply.—

We appreciate the interest in our work. To date, we have received correspondence from the chiropractic community, suggesting that we overstated the data, and the medical community, suggesting that we understated them. The advantage of a systematic review is that it allows for a transparent approach to the subject area to minimize bias. We appreciate the careful review our work has received. In particular, 2 errors were identified with regards to material that was correctly cited in tables but incorrectly cited in text (eg, right text in wrong location). We have notified the journal editors so that these errors may be rectified. We would like to address some of the other concerns raised.

The objectives and methods of this review seem to have been misinterpreted by some. The purpose of our systematic review was to assess the safety of pediatric spinal manipulation, not its efficacy or the safety of other interventions. These topics were appropriately omitted from our article. In addition, our review was not limited to cervical manipulation, as Rosner seemed to indicate. Because we could not locate validated adverse-event categories, we chose categories/definitions that are consistent with those used by the National Institutes of Health. By definition, minor events had to be self-resolving and not require additional medical care. Each of the minor adverse events reported in our article met the a priori definition. We did not think it was reasonable to omit the delayed diagnosis/treatment that was revealed through our systematic review, and we, therefore, applied the same categories to classify the resulting harms.

We actively collaborate with a variety of complementary and alternative medicine institutions, including a national chiropractic college. Our recent research has
Noninitiation or Withdrawal of Intensive Care for High-Risk Newborns

To the Editor.—

The recently published policy statement of the American Academy of Pediatrics Committee on Fetus and Newborn tackles the difficult decisions that parents and health care professionals have to make about the treatment and care of extremely premature or very ill infants.1 Being prepared simultaneously was a report of the Nuffield Council on Bioethics titled “Critical Care Decisions in Fetal and Neonatal Medicine: Ethical Issues,”2 which was published in November 2006. The council’s recommendations on withdrawing and withholding treatment have much in common with those of the American Academy of Pediatrics committee. Both bodies concluded that it is not always right to put an infant through the stress and pain of invasive treatment if he or she is unlikely to improve and death is inevitable, and both recommended that every effort should be made to secure consensus about treatment between the parents and the health care team.

The Nuffield Council on Bioethics also went on to propose, as the basis for professional discussion, week-by-week guidelines to assist decisions on when to initiate intensive care for extremely premature infants. The aim of the guidelines was to help parents and doctors make more informed decisions in a way that would be more open and consistent. They are not intended to be rigid rules, and each case will always need to be considered individually. The guidelines should be reviewed regularly to reflect any changes in outcome for premature infants.

The proposed guidelines include the following:

- Infants born before 22 weeks’ gestation should only be given intensive care as part of a research study.
- Between 22 weeks, 0 days, and 22 weeks, 6 days, standard practice should be to not resuscitate the infant. Resuscitation should only be attempted and intensive care offered if the parents request resuscitation, and reiterate this request, after thorough discussion with an experienced pediatrician about the risks and long-term outcomes and if the clinicians agree that it is in the infant’s best interests.
- Between 23 weeks, 0 days, and 23 weeks, 6 days, it is very difficult to predict the future outcome for an individual infant. Precedence should be given to the wishes of the parents. However, when the condition of the infant indicates that he or she will not survive for long, clinicians should not be obliged to proceed with treatment wholly contrary to their clinical judgment, if they judge that treatment would be futile.
- Between 24 weeks, 0 days, and 24 weeks, 6 days, normal practice should be to provide full invasive intensive care and support from birth and to admit the infant to a NICU unless the parents and clinicians are agreed that the infant’s condition is such that it is not in his or her best interests to start intensive care.
- At ≥25 weeks’ gestation, the relatively high rate of survival and the relatively low risk of severe disability are such that intensive care should be initiated and the infant admitted to a NICU unless he or she is known to be affected by some severe abnormality that is incompatible with any significant period of survival.

Professional bodies in the United Kingdom are currently considering the council’s recommendations. For additional information or to download a copy of the report, see www.nuffieldbioethics.org.

Hugh Whittall, BA
Nuffield Council on Bioethics
London WC1B 3JS, United Kingdom

REFERENCES


In Reply.—

The members of the American Academy of Pediatrics Committee on Fetus and Newborn thank Mr Whittall and the Nuffield Council on Bioethics for their correspondence and for calling attention to the recent Nuffield report, “Critical Care Decisions in Fetal and Neonatal Medicine: Ethical Issues.” As Mr Whittall pointed out, the Nuffield report provides a practical framework to guide decisions based on gestational age for infants born before 25 weeks’ gestation. The American Academy of Pediatrics Neonatal Resuscitation Program (NRP) Steering Committee took a similar approach in its 2005 NRP guidelines.2 They cited the infant born before 23 weeks’ gestation as an example of the category of infants for whom resuscitation would not be indicated, and they
cited the infant of ≥25 weeks’ gestation as an example of one for whom resuscitation is nearly always indicated. For infants of 23 and 24 weeks’ gestation, those with uncertain prognosis, they indicated that parental wishes should take precedence.

The Committee on Fetus and Newborn considered such an approach (ie, defining the gestational age below which resuscitation should not be undertaken and another gestational age above which resuscitation should be routine), with decisions in the intermediate range guided primarily by the wishes of informed parents. However, we decided against this categorical approach for several reasons. First, the prognosis for extremely premature infants is influenced by other factors besides gestational age, such as fetal gender and corticosteroid exposure. Second, we avoided proposing gestational-age guidelines because we felt that these demarcation lines are not fixed but have moved over time and may continue to do so. Third, the concept of standard gestational-age cutoffs for active intervention is problematic because there is considerable center-to-center variability in outcome below 25 weeks’ gestation, both in mortality and morbidity; ideally, decisions should be based on prognosis. However, much of the variability in outcome among centers results from differences in the attitudes and beliefs of medical staff regarding the potential for intact survival at very early gestation. In fact, it is not possible to isolate the impact of the attitudes of medical providers on prognosis from other factors that contribute to center-to-center variability in outcome in the United States and other medically advanced countries.

The worldwide variation in potential outcome at the threshold of viability dwarfs the magnitude of center-to-center variability in the United States and United Kingdom. The guidelines of the Nuffield Council on Bioethics and the NRP Steering Committee may be reasonable for the United States and the United Kingdom in 2007, but these specific gestational-age boundaries would not be appropriate in many parts of the world where maternal-fetal medicine and neonatal intensive care are not as advanced.1)

To be truly useful, an approach using default gestational-age cutoffs to guide therapy would have to allow for consideration of other factors in addition to completed weeks of gestation. As an example, a female fetus at 24 weeks, 6 days, whose mother received betamethasone has a better prognosis than a male fetus at 24 weeks, 0 days, whose mother did not receive corticosteroid. In addition, the guidelines would need to be reappraised every few years as new outcome information becomes available.

The Nuffield report is a thorough and scholarly document. Insofar as treatment decisions at the threshold of viability are sometimes made primarily on the basis of gestational age, the guidelines proposed in the Nuffield report are valid in medically advanced countries in the early 21st century. We commend the Nuffield Council on Bioethics for their efforts in producing this important document, and we welcome the discussion it has stimulated.

**REFERENCES**

Neonatal Blue-Light Phototherapy Could Increase the Risk of Dysplastic Nevus Development

To the Editor.—

We are extremely grateful for the comments by Drs Maisels and Newman1 concerning our report2 on the effect of neonatal blue-light phototherapy on dysplastic nevus development, and we are happy to respond to their questions. Drs Maisels and Newman questioned how the study population had been selected and how the data regarding the neonatal history of the children had been obtained. Our study was conducted in 2 secondary schools on an unselected study population of 14- to 18-year-olds. Each subject underwent a whole-body examination, excluding the scalp and the anogenital area. Before the start of the survey, a standardized diagnostic protocol was defined for the identification of common acquired and clinically atypical melanocytic nevi. After the clinical examination, a standardized questionnaire was completed by all the students in consultation with their parents. The questionnaire data concerning the neonatal history of the children were compared with the official neonatal medical charts to validate our results.

Drs Maisels and Newman also raise the very important question of the light source used for the treatment of neonatal jaundice in Hungary in the 1980s. Blue-light phototherapy has been used in Hungary to reduce the plasma concentration of bilirubin for decades. The spectrum of the blue-light lamp is between 370 and 600 nm, with a maximum at 450 nm. Approximately 0.3% of the emitted light comprises UV-A radiation.

We agree that a notably high proportion of the newborns in Hungary received neonatal phototherapy as compared with those in the United States. It is our opinion that it would be advisable to define a more restricted treatment protocol to rule out the unnecessary application of blue-light phototherapy and thereby prevent the potential adverse effects.

Zsanett Csoma, MD, PhD
Department of Dermatology and Allergology
Peter Hencz, MD, PhD
Department of Pediatrics
Hajnalka Orvos, MD, PhD
Department of Obstetrics and Gynecology
Lajos Kemeny, MD, DSc
Attila Dobozy, MD, DSc
Eva Dosa-Racz
Zsuzsanna Erdei, MD
Dora Bartusek, MD
Judit Olah, MD, PhD
Department of Dermatology and Allergology
University of Szeged
H-6720 Szeged, Hungary

REFERENCES

doi:10.1542/peds.2007-0857
ERRATA


Errors occurred in the article by Patel et al, titled “Association of Proinflammatory Cytokine Gene Polymorphisms With Susceptibility to Otitis Media,” published in the December 2006 issue of Pediatrics (doi:10.1542/peds.2006-0764). In Tables 2 and 3 on pages 2275 and 2276, respectively, the authors reported the genotypes for TNF-α⁻³⁰⁸ and IL-6⁻¹⁷⁴ incorrectly. For TNF-α⁻³⁰⁸, footnote “c” should be assigned to “G/G” and footnote “d” should be assigned to “G/A or A/A.” For IL-6⁻¹⁷⁴, footnote “c” should be assigned to “G/G” and footnote “d” should be assigned to “G/C or C/C.” On page 2277, Discussion section, second paragraph, line 7, “IL-6⁻¹⁷⁴ GG polymorphism” should be replaced with “TNF-α⁻³⁰⁸ AA/AG polymorphism.”

doi:10.1542/peds.2007-1095


doi:10.1542/peds.2007-1030

**Claudius I, Keens T. Do All Infants With Apparent Life-Threatening Events Need to Be Admitted?** PEDIATRICS 2007;119:679–683.

An error occurred in the article by Claudius and Keens, titled “Do All Infants With Apparent Life-Threatening Events Need to Be Admitted?” published in the April 2007 issue of Pediatrics (doi:10.1542/peds.2006-2549). On page 679, in the Results section of the Abstract, on lines 4-6, the authors wrote: “In our study group, the high-risk criteria of age of <1 year and multiple apparent life-threatening events yielded a negative predictive value of 100% to identify the need for hospital admission.” It should read: “In our study group, the high-risk criteria of age of <1 month and multiple apparent life-threatening events yielded a negative predictive value of 100% to identify the need for hospital admission.”

doi:10.1542/peds.2007-1123
Sicherer SH, Simons FER; Section on Allergy and Immunology. Self-injectable Epinephrine for First-Aid Management of Anaphylaxis. PEDIATRICS 2007;119:638–646.

An error occurred in the American Academy of Pediatrics clinical report “Self-injectable Epinephrine for First-Aid Management of Anaphylaxis” published in the March 2007 issue of Pediatrics (doi:10.1542/peds.2006-3689). On page 640, under the heading Epinephrine Autoinjectors: 0.15 or 0.30 mg?, line 10, the authors wrote: “(0.012 mg/kg) rather than an underdose (0.06 mg/kg).” It should read: “(0.012 mg/kg) rather than an underdose (0.006 mg/kg).”

doi:10.1542/peds.2007-1193


An error occurred in the article by Clark et al, titled “A Randomized, Controlled Trial of Acetaminophen, Ibuprofen, and Codeine for Acute Pain Relief in Children With Musculoskeletal Trauma,” published in the March 2007 issue of Pediatrics (doi:10.1542/peds.2006-1347). On page 462, Data Analysis section, lines 8–11, the authors wrote: “Categorical outcomes (such as adequate analgesia achieved) were compared using χ² tests or Fisher’s exact tests when necessary.” It should read: “The number of patients achieving adequate analgesia was stratified for baseline VAS score (below or above 30 mm) and compared using study groups using a McNemar 3-way test. Other categorical outcomes (such as occurrence of adverse events or effects) were compared using χ² tests or Fisher’s exact tests when necessary.”

doi:10.1542/peds.2007-1194


An error occurred in the article by Nord et al, titled “Multiple Cutaneous Infantile Hemangiomas Associated With Hepatic Angiosarcoma: Case Report and Review of the Literature,” published in the September 2006 issue of Pediatrics Electronic Pages (doi:10.1542/peds.2006-0183). In Table 1 on page e911, the authors wrote “Died” as the outcome of case 7. It should read “Alive.”

doi:10.1542/peds.2007-1196

An error occurred in the article by Wolak et al, titled “Wanted and Unwanted Exposure to Online Pornography in a National Sample of Youth Internet Users,” published in the February 2007 issue of Pediatrics (doi:10.1542/peds.2006-1891). In the footnote on page 247, the authors wrote: “To comply with Section 507 of Public Law 104-208 (the Stevens Amendment), we advise readers that 100% of the funds for this research were derived from federal sources through the National Center for Missing and Exploited Children and the US Department of Justice Office of Juvenile Justice and Delinquency Prevention. The total amount of federal funding involved was $348 767.” It should read: “To comply with Section 507 of Public Law 104-208 (the Stevens Amendment), we advise readers that 100% of the funds for this research were derived from federal sources, through grant 2005-MC-CX-K024 from the Office of Juvenile Justice and Delinquency Prevention, US Department of Justice, and grant HSCEOP-05-P-00346 from the Department of Homeland Security, US Secret Service. The total amount of federal funding was $348 767. Points of view or opinions in this article are those of the authors and do not necessarily represent the official position or policies of the US Department of Justice or Department of Homeland Security.”

doi:1031542/peds.2007-1197