

# Are Anesthesia and Surgery during Infancy Associated with Altered Academic Performance during Childhood?

Robert I. Block, Ph.D.,\* Joss J. Thomas, M.D.,† Emine O. Bayman, Ph.D.,‡ James Y. Choi, M.D.,† Karolie K. Kimble, R.N., B.A.,§ Michael M. Todd, M.D.||

## ABSTRACT

**Background:** Although studies in neonatal animals show that anesthetics have neurotoxic effects, relevant human evidence is limited. We examined whether children who had surgery during infancy showed deficits in academic achievement.

**Methods:** We attempted to contact parents of 577 children who, during infancy, had one of three operations typically performed in otherwise healthy children. We compared scores on academic achievement tests with population norms.

**Results:** Composite scores were available for 287 patients. The mean normal curve equivalent score was  $43.0 \pm 22.4$  (mean  $\pm$  SD), lower than the expected normative value of 50,  $P < 0.0001$  by one-sample Student  $t$  test; and 35 (12%) had scores below the 5th percentile, more than expected,  $P < 0.00001$  by binomial test. Of 133 patients who consented to participate so that their scores could be examined in relation to their medical records, the mean score was  $45.9 \pm 22.9$ ,  $P = 0.0411$ ; and 15 (11%) scored below the 5th percentile,  $P = 0.0039$ . Of 58 patients whose medical records showed no central nervous system problems/potential risk factors during infancy, 8 (14%) scored below the 5th percentile,  $P = 0.008$ ; however, the mean score,  $47.6 \pm 23.4$ , was not significantly lower than expected,  $P = 0.441$ . Duration of anesthesia and surgery correlated negatively with scores ( $r = -0.34$ ,  $N = 58$ ,  $P = 0.0101$ ).

\* Associate Professor, † Clinical Associate Professor, § Research Assistant, || Professor and Head of Department, Department of Anesthesia, Roy J. and Lucille A. Carver College of Medicine, University of Iowa, Iowa City, Iowa. ‡ Assistant Professor, Department of Anesthesia, Roy J. and Lucille A. Carver College of Medicine, University of Iowa, and Department of Biostatistics, College of Public Health, University of Iowa.

Received from the Department of Anesthesia, Roy J. and Lucille A. Carver College of Medicine, University of Iowa, Iowa City, Iowa. Submitted for publication July 11, 2011. Accepted for publication May 11, 2012. Support was provided solely from institutional and/or departmental sources. Presented at the annual meeting of the International Anesthesia Research Society, International Anesthesia Research Society and SAFEKIDS International Science Symposium: Anesthetic-Induced Neonatal Neuronal Injury, March 20, 2010, Honolulu, Hawaii.

Address correspondence to Dr. Block: Anesthesia Research, Westlawn Building, Room 5140, University of Iowa, Iowa City, Iowa 52242. robert-block@uiowa.edu. This article may be accessed for personal use at no charge through the Journal Web site, www.anesthesiology.org.

Copyright © 2012, the American Society of Anesthesiologists, Inc. Lippincott Williams & Wilkins. Anesthesiology 2012; 117:494–503

## What We Already Know about This Topic

- Some studies have demonstrated an association between anesthesia and surgery early in life and developmental delays or deficits
- Previous studies are limited by examining infants with other potential causes for delay

## What This Article Tells Us That Is New

- In 58 infants without other risk factors for central nervous system problems who underwent inguinal hernia repair and orchiopexy, pyloromyotomy, or circumcision under general anesthesia, mean standardized achievement test scores at school age did not differ from normal
- There was an association between test score and duration of anesthesia exposure among these children, and a greater-than-expected number of them had very low test performance (less than 5th percentile)

**Conclusions:** Although the findings are consistent with possible adverse effects of anesthesia and surgery during infancy on subsequent academic achievement, other explanations are possible and further investigations are needed.

ALTHOUGH studies in late fetal and neonatal animals show that anesthetics have neurotoxic effects,<sup>1,2</sup> relevant human evidence is limited. Few well-designed studies have evaluated possible long-term adverse neurologic outcomes of anesthesia and surgery in young children. Some clinical studies examined such outcomes, but many involved patients with significant central nervous system (CNS) problems or potential risk factors for such problems.<sup>1,3</sup> Such patients may be at risk for long-term neurodevelopmental delays and may be difficult to match with unexposed controls, increasing the likelihood of confounding. For example, extremely premature or low-birth-weight infants exposed to anesthesia and surgery or prolonged sedation and/or analgesia have lower gestational ages and birth weights than unexposed infants.<sup>4–6</sup>

Recent studies focusing on largely “normal” populations have produced mixed results. Apart from one pilot study, which found no significant differences in the incidence of

◆ This article is accompanied by an Editorial View. Please see: Flick RP, Warner DO: A users' guide to interpreting observational studies of pediatric anesthetic neurotoxicity: The lessons of Sir Bradford Hill. *ANESTHESIOLOGY* 2012; 117:459–62.

behavioral disturbances in children who were exposed to anesthesia for urological procedures at younger *versus* older ages,<sup>7</sup> all were entirely retrospective. Three studies, which included children undergoing any surgery, assessed the influence of anesthesia (and surgery) on subsequent learning disabilities, academic achievement and cognitive test performance, or individualized education programs.<sup>8–10</sup> These studies found some deficits associated with anesthesia and surgery, but in two studies they only occurred with multiple operations,<sup>8,9</sup> and in the other, a twin study, they were interpreted as not caused by anesthesia.<sup>10</sup> Another study, which assessed the influence of inguinal hernia repair on subsequent academic achievement test performance and teacher ratings, found no deficit among children for whom test scores were available.<sup>11</sup> In the first of two studies by DiMaggio *et al.*, hernia repair was associated with an increased frequency of subsequent developmental or behavioral disorders.<sup>12</sup> Their second study showed generally similar results in overall analyses for twins who had varied surgeries.<sup>13</sup>

The present study of effects of anesthesia and surgery on subsequent academic achievement combined four features that were not jointly present in any previous study and sought to increase the likelihood of detecting effects or help clarify the nature of any effects. First, focusing on effects of surgery during infancy, based on the assumption that earlier ages of surgery might be associated with greater long-term effects.<sup>14–16</sup> Next, studying three groups of operations that are frequently performed in otherwise healthy, normal infants. Doing so allowed comparisons to examine the possibility that infants requiring one particular type of operation might have preexisting CNS or cognitive impairments that could confound effects of anesthesia and surgery on academic achievement; no such examination would be feasible by studying one or a wide variety of operations. Third, the study included focusing on a subgroup of patients who did not have identifiable CNS problems/potential risk factors during infancy, which might confound effects of anesthesia and surgery. Lastly, the study examined the relationship of academic achievement to duration of anesthesia and surgery in this subgroup.

## Materials and Methods

### Approval

The research and consent procedures were approved by the University of Iowa Institutional Review Board (Iowa City, Iowa).

### Selection of Groups of Operations

Three groups of operations were selected for the research through a pilot study that attempted to identify operations that were often performed on otherwise healthy infants (*i.e.*, age younger than 1 yr) and for which sufficient numbers of cases were available for review: inguinal hernia repair and orchiopexy (with or without hernia repair), pyloromyotomy, and circumcision performed under general anesthesia.

### Selection of Patients

Department of Anesthesia billing records were searched for patients who had one or more of the three groups of operations during infancy and who were between 7.0 and 17.9 yr old on the date of the search (January 28, 2008). A total of 623 patients were initially identified. Forty-six were eliminated because the patients had subsequently died, were determined not to have received general anesthesia, or for other reasons (table 1).

### Mailings to Parents

To authorize retrieval of academic achievement test scores and to allow us to compare these scores with information in the medical record, parents had to provide written informed consent and patients themselves had to provide written informed consent (ages 13.0 yr or older) or assent (ages below 13.0 yr). Mailings were sent to parents of 577 patients. Brief initial letters inviting participation in the study were sent, followed shortly thereafter by packets of materials including informed consent/assent documents and forms for authorizing retrieval of test scores and detailing the history of schools attended by the patient. Optionally, parents could also agree to participate in a brief telephone interview. A follow-up packet was sent to parents who did not respond. Patients and parents were compensated for participation.

Our hospital's mailing addresses for parents were frequently out of date because of the length of time between the date of surgery and the present. Current mailing addresses and telephone numbers were sought through Internet searches and the MetroNet database of Experian, a credit reporting agency. Efforts were made to contact parents by telephone when mailing addresses were uncertain. To aid in tracking patients, Iowa Testing Programs of the University of Iowa, which administers the statewide achievement testing program in Iowa, provided some information concerning school districts in which score records could be located.

### Retrieval of Academic Achievement Test Scores

We retrieved scores on the Iowa Tests of Basic Skills and Iowa Tests of Educational Development (hereafter, "Iowa Tests"). The Iowa Tests are widely used, standardized tests assessing basic, general intellectual skills and abilities in verbal, mathematical, and other areas.<sup>17</sup> Scores were provided by Iowa Testing Programs from their database of testing within Iowa. When consent was given, records of patients' scores, identified by name, were provided. When consent was not obtained, but patients did not specifically decline to participate, deidentified records of patients' scores were provided when they could be located. (In this subgroup, it was not possible to directly link test scores with medical record information.)

For analysis, we used the complete Iowa Test composite score in the earliest grade for which this score was available (for 89% of patients, this was in grades two through four, at which time they were 7–10 yr old). This score summarized

**Table 1.** Recruitment of Patients and Retrieval of Iowa Test Scores

Status of Patients	N
Patients identified, surgery at younger than 1 yr, current age 7.0–17.9 yr	623
Patients eliminated*	46
Packets inviting participation sent	577
Returned as undeliverable, not tracked**	93
Not returned as undeliverable, no contact**	215
Some contact	269
Agreed to participate, returned forms	231
Iowa Test scores retrieved	158
No Iowa Test scores available, exclusively attended out of state schools	42
No Iowa Test scores available, other reasons***	27
Forms returned incompletely and never returned completed forms	4
Phone contact only	38
No Iowa Test scores available, exclusively attended out of state schools	16
Not interested in participating for various reasons (e.g., home schooled, incarcerated mother)	22

Indentations in the table represent subcategories of the category above.

The mailings to 577 parents of patients resulted in 231 consents to participate in the research, a 40% participation rate. Only a small percentage of those to whom packets were mailed ( $22/577 = 4\%$ ) explicitly refused to participate. Much more commonly, we could not find current contact information ( $93/577 = 16\%$ ) or received no response to the mailings ( $215/577 = 37\%$ ). Some Iowa Test scores (*i.e.*, subtest, but not necessarily composite, scores) were retrieved for  $158/231 = 68\%$  of patients who consented to participate.

\* Eighteen were dead; four were determined not to have received general anesthesia; and 12 were determined not to have had the surgery indicated in the Department of Anesthesia billing records, based on review of their hospital medical records. For 12, there was uncertainty about whether they had had the surgery indicated in the Department of Anesthesia billing records, based on review of their hospital electronic medical records, and this had not yet been resolved by review of their hospital paper medical records at the time of the mailings.

\*\* Iowa Testing Programs located and provided deidentified Iowa Test composite scores for 154 of these 208 patients whose current addresses could not be determined or who did not respond to our mailings.

\*\*\* In almost all cases, these were younger children who attended school in districts that chose not to administer the Iowa Tests until grades 3 or 4, which these children had not yet reached; or children in such districts who had moved out of state prior to those grades.

performance on tests of reading, language, mathematics, science, and social studies, and provided the most comprehensive indicator of overall academic achievement. The composite scores were expressed as percentile scores based on population data for the state of Iowa.

Some of the test scores retrieved for the present study were based on a revised form of the tests that was introduced during the 2001–2002 school year and some were based on the previous form. The revised form was equated to the previous form by administering them jointly to an equating sample.<sup>17</sup> The percentile scores that we analyzed are comparable across grades and represent each student's performance relative to the performance of students in the same grade on the same form of the test administered during a standardization process. Therefore, the percentile scores based on the earlier and revised forms could appropriately be pooled for analysis.

### Telephone Interview

A research assistant contacted parents who consented to the optional telephone interview. This interview covered demographic information (race/ethnicity; siblings; and parental income, education, and employment) and interest in participation in possible future research. In a small number of cases ( $N = 16$ ), when telephone contact was not feasible, the information was provided by mail.

### Medical Record Review

Of the 577 patients to whose parents packets were sent, medical records were not relevant for 58 patients who never attended school in Iowa and for whom no test results were available. The complete medical records at our hospital during patients' infancy were reviewed for the other 519 patients. Information was extracted in prespecified formats in the following categories: demographics, birth, details of the operations from the three groups of operations selected for study, other exposures to anesthetics, and 18 types of prespecified conditions or procedures defined as CNS disorders or potential risk factors for subsequent developmental or cognitive dysfunction (table 2). In addition, information was recorded concerning other brain diseases or conditions and other diagnoses (*e.g.*, respiratory and circulatory conditions) that might potentially be associated with neurologic deficits. Patients were classified as having definable CNS problems/potential risk factors during infancy if they had any of the 18 prespecified conditions or procedures or other CNS problems/potential risk factors. Two pediatric anesthesiologists made this classification independently. Any discrepancies were resolved by discussion. Certain additional details, *e.g.*, intraoperative problems, concerning the three groups of operations selected for study and other exposures to anesthetics were recorded for the 185 patients who agreed to participate, returned the forms completely, and attended some school in Iowa.

**Table 2.** Prespecified Conditions or Procedures Defined as Central Nervous System Problems/Potential Risk Factors

Condition or Procedure
Acute respiratory distress syndrome
Birth asphyxia
Cleft lip
Cleft palate
Central nervous system infections
Congenital anomalies or deformities of skull, face, or jaw
Congenital syndromes associated with mental retardation (e.g., Down syndrome)
Extracorporeal membrane oxygenation
Hydrocephalus
Intracranial hemorrhage
Low birth weight (less than 2,500 g)
Meningomyelocele
Microcephaly
Periventricular leukomalacia
Prematurity (less than 37 weeks gestational age at birth)
Seizure/convulsive disorder (not febrile-related)
Stroke
Surgery requiring cardiopulmonary bypass

### Statistical Analysis

Using the composite scores expressed as percentile scores described above, the percentages of patients who performed very poorly (*i.e.*, below the 5th percentile) were compared, using binomial tests, with the Iowa population, *i.e.*, all children in Iowa taking the Iowa Tests. Proportions of composite scores below the 5th percentile were compared between subgroups using Fisher exact tests. A significance level of  $P < 0.05$  with two-tailed tests was used in all analyses. Values of  $P \geq 0.05$  indicate nonsignificant standard hypothesis testing results. These are described as nonsignificant differences in the Results; we did not test for, or demonstrate, equivalence of values. Statistical analyses were done using SAS 9.2 (SAS Institute Inc., Cary, NC) and R 2.11.1 (R Foundation for Statistical Computing, Vienna, Austria).

For parametric analyses, the percentile scores (which are uniformly rather than normally distributed) were converted to normal curve equivalent scores. Normal curve equivalent scores, which are commonly used in educational testing, have a range similar to percentile scores (1–99) but are normally distributed. They were compared with the expected value of 50 for a normative population with one-sample Student *t* tests and between subgroups with two-sample Student *t* tests.

A subsequent, more detailed analysis was based on a subgroup of patients without CNS problems/potential risk factors during infancy who consented to participate and hence for whom patient-identified composite scores could be retrieved. The empirical cumulative distribution function of percentile scores was examined. The relationship of normal curve equivalent scores to duration of anesthesia and surgery was analyzed using linear regression.

**Table 3.** Most Common Central Nervous System Problems/Potential Risk Factors

Central Nervous System Problem/Potential Risk Factor	Percentage of Patients
Prematurity	73
Low birth weight	62
Acute respiratory distress syndrome	53
Intracranial hemorrhage	9
Congenital anomalies or deformities of skull, face, or jaw	8

Percentages are based on the 75 patients with Iowa Test composite scores who had at least one of the central nervous system problems/potential risk factors listed in table 1.

Differences among groups of operations were compared by one-way ANOVA for quantitative characteristics and Fisher exact test for categorical characteristics.

## Results

### Patients Available for Study and Outcomes

The numbers of patients available for the study and our results in recruiting them for the research and retrieving Iowa Test scores are summarized in table 1. Of the 158 patients who consented to participate and for whom some identified Iowa Test scores (*i.e.*, subtest, but not necessarily composite, scores) were retrieved, a composite score was available for  $133/158 = 84\%$ . Of the patients with composite scores,  $58/133 = 44\%$  were classified as having no CNS problems/potential risk factors and  $75/133 = 56\%$  as having at least one CNS problem/potential risk factor. The most common CNS problems/potential risk factors are shown in table 3. In addition, Iowa Testing Programs located and provided de-identified composite scores for an additional 154 patients whose current addresses could not be determined or who did not respond to our mailings (and whose status with respect to CNS problems/potential risk factors, was, therefore, unknown).

### Average Test Scores

For the 287 patients with available composite scores, the mean normal curve equivalent score was  $43.0 \pm 22.4$  (mean  $\pm$  SD), significantly lower than the expected value of 50 for a normative population,  $P < 0.0001$  by one-sample Student *t* test, 95% CI = 40.4, 45.6. Of 133 consented patients with composite scores, the mean normal curve equivalent score was  $45.9 \pm 22.9$ , significantly lower than expected,  $P = 0.0411$ , 95% CI = 42.0, 49.8. Of the 58 of these 133 patients whose medical records showed no CNS problems/potential risk factors during infancy, the mean normal curve equivalent score,  $47.6 \pm 23.4$ , was not significantly lower than expected,  $P = 0.441$ , 95% CI = 41.4, 53.8.

### Patients with Very Poor Academic Achievement

Of the 287 patients with available composite scores, 35 (12%) had Iowa Test scores below the 5th percentile (95% CI = 8.6%, 16.6%), compared with 5% of the Iowa population,  $P < 0.00001$  by binomial test. Of the 133 consented patients with composite scores, 15 (11%) scored below the 5th percentile (95% CI = 6.5%, 17.9%), more than in the Iowa population,  $P = 0.0039$ . Of the 58 of these 133 patients who had no CNS problems/potential risk factors, 8 (14%) scored below the 5th percentile (95% CI = 6.1%, 25.4%), more than in the Iowa population,  $P = 0.008$ .

### Subgroup Comparisons

The percentages of patients who scored below the 5th percentile did not differ significantly between patients with and without CNS problems/potential risk factors, 9% (7/75) and 14% (8/58), respectively (Fisher exact test,  $P = 0.582$ , odds ratio [OR] = 0.643, 95% CI = 0.219, 1.891); or between patients who consented to participate and nonconsented patients whose deidentified composite scores were obtained, 11% (15/133) and 13% (20/154) (Fisher exact test,  $P = 0.720$ , OR = 1.174, 95% CI = 0.575, 2.397). The latter finding suggests that there was no evident bias toward children with very poor academic performance being either more or less inclined to consent to participate. The mean normal curve equivalent scores did not differ significantly between patients with and without CNS problems/potential risk factors ( $44.6 \pm 22.5$  and  $47.6 \pm 23.4$ , respectively, Student  $t$  test not significant,  $P = 0.453$ , mean difference = 3.0, 95% CI = -4.9, 10.9), but were higher for patients who consented to participate than nonconsented patients whose deidentified composite scores were obtained ( $45.9 \pm 22.9$  and  $40.4 \pm 21.8$ ,  $P = 0.0384$  by Student  $t$  test, mean difference = 5.5, 95% CI = 0.3, 10.7).

### Patients with No CNS Problems or Potential Risk Factors

More detailed analyses of composite scores were limited to the 58 patients without CNS problems/potential risk factors, because such problems/potential risk factors might confound effects of anesthesia and surgery. These 58 patients were 84% male; 3% were Hispanic and 97% were non-Hispanic. Their race was 91% white, 2% black, and 7% "other." The drugs used for anesthesia for these patients, and the percentages of patients who received them, are shown in table 4. Data concerning income, education, and employment of the patients' parents were obtained during the telephone interview and were similar to statewide data for Iowa (data not shown).

One patient had both orchiopexy and circumcision on the same date; the remainder had only one of the selected groups of operations (inguinal hernia repair/orchiopexy, pyloromyotomy, and circumcision for  $N = 17$ , 24, and 16, respectively). Five patients (9%) had additional procedures on the same dates, and four patients (7%) had additional

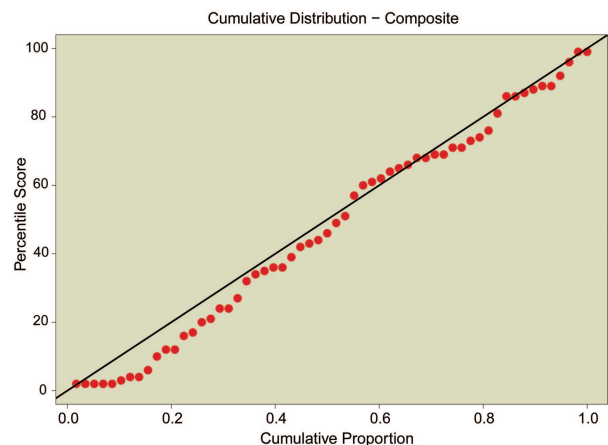
**Table 4.** Drugs Used for Anesthesia for the 58 Patients Without Central Nervous System Problems/Potential Risk Factors

Drug	Percentage of Patients
Nitrous oxide	72
Halothane	71
Thiopental	24
Isoflurane	22
Morphine	22
Sevoflurane	12
Fentanyl	10
Desflurane	2
Ketamine	2
Propofol	2

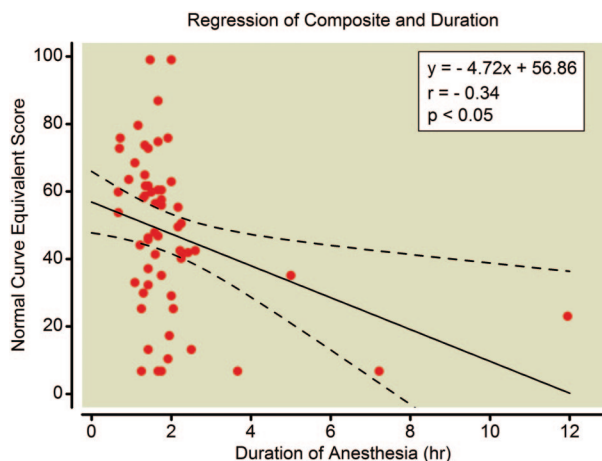
operations outside the selected groups of operations on additional dates during infancy. There was little information in the patients' medical records indicating exposure to anesthesia or sedation during birth (5%) or *in utero* (0%).

### Distribution of Composite Scores

The empirical cumulative distribution function for the Iowa Test composite scores of the 58 patients without CNS problems/potential risk factors is shown by the circles in figure 1. For any specific percentile score, the extent (if any) to which the circle representing patients lies below the straight line indicates the extent to which patients have poorer scores than the Iowa population. The circles representing patients were substantially below the straight line at the lower percentiles and did not cross it until the 57<sup>th</sup> percentile. The overrepresentation of patients scoring below the 5th percentile is evident in the figure. However, this overrepresentation diminished at the level of median performance, *i.e.*, 52% of patients (30/58) scored below the 50th percentile, compared with 50% of the Iowa population.



**Fig. 1.** Empirical cumulative distribution function for the Iowa Test composite scores of the 58 patients without central nervous system problems/potential risk factors. The circles indicate the proportions of patients scoring at or below each of the percentile scores. The straight line represents the distribution for the Iowa population, *i.e.*, all children in Iowa taking the Iowa Tests.



**Fig. 2.** Relationship of Iowa Test normal curve equivalent scores to durations of anesthesia (which correlate highly with durations of surgery) for the 58 patients without central nervous system problems/potential risk factors. In addition to the durations for the selected groups of operations (inguinal hernia repair/orchiopexy, pyloromyotomy, and circumcision), the plotted values include the durations for the four patients mentioned in the text for the additional operations outside the selected groups of operations on additional dates during infancy. The *solid line* is the regression line and the *dashed lines* indicate 95% CI.

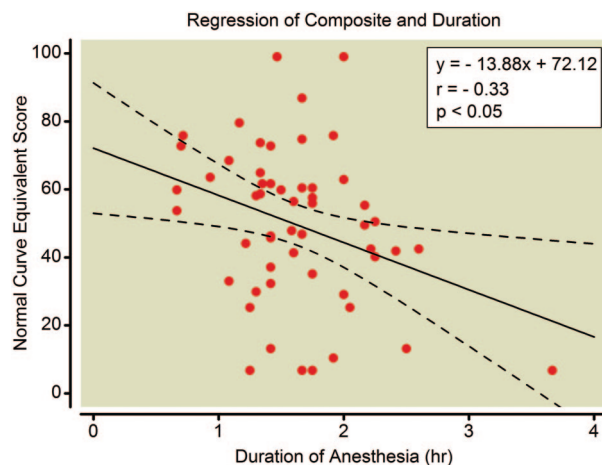
### Patients with a Single Operation during Infancy

The percentage of patients who scored below the 5th percentile was also examined after excluding the four patients mentioned above who had additional operations on additional dates during infancy outside the selected groups of operations, leaving only the patients who had a single operation date during infancy. In this subgroup, 13% (7/54) scored below the 5th percentile (95% CI = 5.4%, 24.9%), compared with 5% of the Iowa population,  $P = 0.0176$ .

### Composite Scores and Durations of Anesthesia and Surgery

Figures 2 and 3 show the relationships of normal curve equivalent scores to durations of anesthesia (which correlate highly with the durations of surgery). Both figures include the durations for the selected groups of operations. Figure 2 includes, in addition, the durations for the four patients mentioned above for the additional operations outside the selected groups of operations on additional dates during infancy. Figure 3, in contrast, excludes these four patients and shows data for the remaining 54 patients, who each had only a single operation date during infancy.

As shown in figure 2, normal curve equivalent scores were lower in patients with longer durations relative to those with shorter durations. The Pearson product moment correlation was  $r = -0.34$ ,  $P = 0.0101$ , 95% CI =  $-0.55, -0.08$ . When only the patients who had a single operation date during infancy were considered (fig. 3), normal curve equivalent scores were also lower in patients with longer durations,  $r = -0.33$ ,  $P = 0.0165$ , 95% CI =  $-0.55,$



**Fig. 3.** Relationship of Iowa Test normal curve equivalent scores to durations of anesthesia for 54 of the 58 patients without central nervous system problems/potential risk factors. The plotted values are the same as in fig. 2, except that the four patients mentioned in the text who had additional operations outside the selected groups of operations on additional dates during infancy are omitted. Thus, the patients whose values are plotted in fig. 3 each had only a single operation date during infancy. The *solid line* is the regression line and the *dashed lines* indicate 95% CI.

$-0.06$ . Figure 3, compared with figure 2, illustrates an almost identical correlation, but a smaller range of durations and larger magnitudes of both the slope and intercept in the regression equation.

When the correlations were done separately for each of the three groups of operations selected for study, they were consistently negative, but varied in magnitude (not surprisingly, considering the small Ns for these subgroups). For example, those corresponding to figure 2 were  $r = -0.40$  for inguinal hernia repair/orchiopexy,  $-0.20$  for pyloromyotomy, and  $-0.46$  for circumcision.

### Supplemental Analyses

Supplemental analyses were done to examine the representativeness of the results of the primary analyses described above and aid in their interpretation.

Characteristics extracted from our hospital's medical records were compared between patients with whose parents we did or did not establish some contact by U.S. mail, telephone, or electronic mail. Patients who had never attended any school in Iowa were excluded from this comparison, because, even if they had consented to participate, no Iowa Test scores could have been retrieved for them. Contacted and noncontacted patients did not differ significantly in prevalence of CNS problems/potential risk factors or percentages of males or Hispanics (data not shown). There were higher percentages of patients with white race among those contacted than noncontacted patients (90% *vs.* 81%,  $P = 0.008$ , OR = 2.055, 95% CI = 1.204, 3.507).

Mean normal curve equivalent scores did not differ significantly among the three groups of operations ( $45.4 \pm 24.4$

for circumcision,  $46.3 \pm 24.7$  for inguinal hernia repair/orchiopexy, and  $46.3 \pm 18.1$  for pyloromyotomy,  $P = 0.984$  by ANOVA). The percentages of patients who scored below the 5th percentile also did not differ significantly among groups of operations, either for all patients or for those with or without CNS problems/potential risk factors considered separately (data not shown). The percentages of males or Hispanics and prevalence of intraoperative problems did not differ significantly among groups of operations, but prevalence of CNS problems/potential risk factors, age at surgery, and race did differ (data not shown).

Five patients with and three without CNS problems/potential risk factors experienced intraoperative problems. None scored below the 5th percentile.

We defined a *post hoc* subset of patients with “more pronounced” CNS problems/potential risk factors by excluding from the patients with CNS problems/potential risk factors those who were so classified solely because of prematurity that was not extreme (*i.e.*, not less than 26 weeks gestational age), low birth weight that was not extreme (*i.e.*, not less than 1,000 g), and/or acute respiratory distress syndrome. Of the 75 patients with composite scores who had CNS problems/potential risk factors based on our *a priori* criteria, 39 were classified as having “more pronounced” CNS problems/potential risk factors and 36 were not. Comparing the patients with “more pronounced” CNS problems/potential risk factors with the other patients (*i.e.*, these 36 patients plus the 58 without CNS problems/potential risk factors), the mean normal curve equivalent scores were lower for patients with “more pronounced” CNS problems/potential risk factors:  $38.1 \pm 19.6$  and  $49.1 \pm 23.4$ , respectively,  $P = 0.0107$  by Student *t* test, mean difference = 11.0, 95% CI = 2.6, 19.5.

## Discussion

A disproportionate number of children undergoing anesthesia and surgery during infancy had very low achievement test scores (below the 5th percentile), both in our overall sample and the subgroup of 58 patients without CNS problems/potential risk factors. There was also a significant association between scores and duration of anesthesia and surgery in this subgroup. Mean scores were reduced for our overall sample and all patients who consented to participate, but not this subgroup.

### Effects of Anesthesia versus Surgery

This study cannot distinguish among effects of anesthesia, surgery, or their combination. The relationship between scores and duration of anesthesia and surgery is consistent with, but does not strongly support, an influence of anesthesia *per se*. Duration of anesthesia is determined by duration of surgery and may reflect the severity of the underlying problem and extent of surgery.

### CNS Problems/Potential Risk Factors

We focused on patients without CNS problems/potential risk factors, which we assumed might be associated with lower scores, thereby confounding effects of anesthesia and surgery. In fact, CNS problems/potential risk factors did not significantly affect scores, perhaps because of our limited sample size and broad definition of CNS problems/potential risk factors, which included being born 1 day premature. Less frequent, more severe CNS problems/potential risk factors might have clearer effects. Indeed, “more pronounced” CNS problems/potential risk factors were associated with lower mean scores. Some previous studies support this interpretation, *e.g.*, impaired academic achievement is not apparent in otherwise healthy late-preterm infants.<sup>18</sup>

### Effects on Very Poor Academic Achievement versus Mean Scores

Among patients without CNS problems/potential risk factors, anesthesia and surgery were associated with an overrepresentation of very low scores, but there was only weak evidence of any reduction in mean scores. This could occur if anesthesia and surgery had a relatively large adverse effect on a small percentage of patients.

The reduction in mean scores, as well as the overrepresentation of very low scores, were significant in our overall sample and all consented patients (including those with CNS problems/potential risk factors). The interpretation of these mixed findings is unclear. As noted above, we cannot exclude subtle effects of CNS problems/potential risk factors. Moreover, mean scores were lower for nonconsented patients whose deidentified scores were obtained than consented patients, indicating a selection bias. Perhaps some lower-performing children were harder to contact or decided against participation. Had we recruited more of them, mean scores among consented patients, possibly including those without CNS problems/potential risk factors, might have been lower.

However, the proportions of patients with very low scores were fairly uniform in the overall sample and subgroups; there was no detectable difference between nonconsented patients whose deidentified scores were obtained and consented patients.

### Representativeness of the Results

Four issues concerning the representativeness of our primary results should be considered:

(A) Groups of operations. Although some characteristics varied significantly among the three operations studied, we found no evidence for a difference among operations in scores, suggesting that our primary results generalized across these operations. The group sizes were small, however, and no operations of high physiologic complexity were studied.

(B) Gender. Patients without CNS problems/potential risk factors for whom we obtained composite scores were 84% male, *versus* 51% of the same age range of the Iowa

population, our comparison group.<sup>#</sup> Although females generally perform better on language tests, composite scores, which also include performance on mathematics, science, and social studies tests, show only a small mean female advantage.<sup>17</sup> We estimated the female advantage for a sample of composite scores from the Iowa population<sup>17</sup> comparable in distribution by grade in school with the scores of the patients analyzed. This estimate was roughly 0.05 SD. The reduction in mean scores in our overall sample, relative to the population, was far larger, *i.e.*, 0.31 SD. Gender, therefore, seems unlikely to account for all the observed effects, although it may have played some role. There are some reports that females have better outcomes than males following traumatic brain injury.<sup>19</sup>

(C) Other demographic factors. Racial and ethnic characteristics of patients without CNS problems/potential risk factors for whom we obtained composite scores were similar to those of the Iowa population (94% white, 3% black; 4% Hispanic, 96% non-Hispanic), as were their parents' socioeconomic status.<sup>\*\*</sup> There may have been differences in other, unknown demographic characteristics.

(D) Characteristics of contacted and noncontacted patients. Characteristics of contacted and noncontacted patients were similar, the only meaningful difference being more white patients among the former. Because blacks perform worse than whites on the Iowa Tests,<sup>17</sup> this selection bias might have contributed to the higher mean scores for patients who consented to participate than nonconsented patients whose deidentified scores were obtained. Inclusion of more noncontacted patients, and, therefore, more blacks, as participants would be unlikely to have weakened our finding that anesthesia and surgery were associated with more very low scores.

### Limitations

This study cannot exclude effects of potential confounding variables or demonstrate causal relationships of scores to anesthesia and surgery. The study had limitations. It was an observational study, not a controlled trial. Success in contacting patients and obtaining consent to link their scores with their medical records was limited (see table 1), with resulting potential selection biases. We compared patients with the Iowa population rather than a matched control group. We focused on operations and CNS problems/potential risk factors during infancy, not later ages. We focused on academic achievement, the sensitivity of which to the damage seen in animal models<sup>1,2</sup> is unclear.

Conceivably, some patients were misclassified as "without CNS problems/potential risk factors" because we relied exclusively on medical records, which could have been incomplete; and this might have influenced our findings. However,

our observation that CNS problems/potential risk factors did not have marked effects on scores mitigates this concern. Moreover, the Iowa population includes many children with CNS problems/potential risk factors as broadly defined in our study: The Iowa Tests are the primary instrument in Iowa for assessing academic proficiency in grades 3–8 and 11 under the federal No Child Left Behind Act and are administered to 99.3% of all assessed students (personal communication, Tom Deeter, Ph.D., Lead Consultant, Iowa Department of Education, Des Moines, IA, July 27, 2010). Nationally, the frequencies of prematurity and low birth weight were 12.8% and 8.3%, respectively, in 2006.<sup>20</sup> Therefore, most premature or low-birth-weight children in Iowa took the Iowa Tests. A comparison group with fewer CNS problems/potential risk factors would, if anything, have likely strengthened our finding that anesthesia and surgery were associated with more very low scores.

### Comparison of Current Findings to Those of Previous Studies

Two studies examined the effects of exposures to anesthesia and surgery before ages 2<sup>9</sup> or 4<sup>7</sup> on subsequent learning disabilities,<sup>8,9</sup> achievement and cognitive test performance,<sup>9</sup> or individualized education programs.<sup>9</sup> Multiple exposures were associated with some adverse effects on these outcomes.<sup>8,9</sup> The risk for learning disabilities increased with longer durations of exposures.<sup>7</sup> We also found an effect of duration. However, we found effects of a single exposure – an important difference, because, in these previous studies (as in ours), large majorities of exposed children were exposed only once.<sup>8,9</sup>

A Danish study examined effects of inguinal hernia repair during infancy on achievement test scores and teacher ratings at ages 15 and 16.<sup>11</sup> Children with scores available showed no effect of exposure. However, scores were unavailable for more exposed than unexposed children (21% and 13%, respectively), and unavailability was most often because of special needs. Our analyses were based on scores at younger ages, which might have contributed toward our different results. The Iowa testing program may possibly include a greater percentage of children with significant CNS problems/potential risk factors during infancy (see the 99.3% figure in Limitations) than the Danish testing program. If so, mean scores may have been relatively less sensitive in the Danish study than in our analyses including all children.

A twin study<sup>10</sup> investigated possible genetic contributions to effects of anesthesia before ages 3 or 12 on achievement test performance. Some deficits associated with anesthesia were found, but interpreted as not caused by anesthesia. This study relied on parents' survey responses, with no independent verification of exposure to anesthesia or surgery. A critique<sup>21</sup> offered several other reasons why the authors' interpretation "should be qualified due to limitations of the study design and the data."

<sup>#</sup> <http://data.iowadatacenter.org/browse/estimates.html>. Accessed February 16, 2011.

<sup>\*\*</sup> [http://factfinder.census.gov/servlet/DatasetMainPageServlet?\\_program=ACS](http://factfinder.census.gov/servlet/DatasetMainPageServlet?_program=ACS). Accessed February 16, 2011.

Concerns about potentially confounding CNS problems/potential risk factors limit inferences that can be drawn from other relevant data.<sup>22–29</sup>

## Conclusion

Our results suggest that even a single exposure to anesthesia and surgery during infancy is associated with an overrepresentation of very low test scores. Whether this association is causative cannot be determined from our study. However, the association between longer durations of anesthesia and surgery and decreasing scores is suggestive, albeit inconclusive. Our study design mandates caution in interpretation and the findings should be considered tentative until further verification. Larger studies including matched control groups and prospective, randomized trials including comparisons of general *versus* regional anesthesia are essential. Future studies should analyze medical records, including duration of anesthesia and surgery and influences of CNS problems/potential risk factors.

The authors thank Stephen Dunbar, Ph.D., Hieronymus-Feldt Professor of Educational Measurement; Catherine Welch, Ph.D., Professor; and Matthew Whittaker, Ph.D., Research Assistant, Iowa Testing Programs, College of Education, University of Iowa, Iowa City, Iowa, for their advice concerning the Iowa Tests of Basic Skills and Iowa Tests of Educational Development and retrieval of test scores for the research. The authors also thank Julie Weeks, M.P.T., Program Associate, Department of Anesthesia, Roy J. and Lucille A. Carver College of Medicine, University of Iowa, for assistance with study design during initial stages of the research.

## References

- Loepke AW, Soriano SG: An assessment of the effects of general anesthetics on developing brain structure and neurocognitive function. *Anesth Analg* 2008; 106:1681–707
- Mellon RD, Simone AF, Rappaport BA: Use of anesthetic agents in neonates and young children. *Anesth Analg* 2007; 104:509–20
- Patel P, Sun L: Update on neonatal anesthetic neurotoxicity: Insight into molecular mechanisms and relevance to humans. *ANESTHESIOLOGY* 2009; 110:703–8
- Kabra NS, Schmidt B, Roberts RS, Doyle LW, Papile L, Fanaroff A, Trial of Indomethacin Prophylaxis in Preterms Investigators: Neurosensory impairment after surgical closure of patent ductus arteriosus in extremely low birth weight infants: Results from the Trial of Indomethacin Prophylaxis in Preterms. *J Pediatr* 2007; 150:229–34, 234.e1
- Rozé JC, Denizot S, Carbajal R, Ancel PY, Kaminski M, Arnaud C, Truffert P, Marret S, Matis J, Thiriez G, Cambonie G, André M, Larroque B, Bréart G: Prolonged sedation and/or analgesia and 5-year neurodevelopment outcome in very preterm infants: Results from the EPIPAGE cohort. *Arch Pediatr Adolesc Med* 2008; 162:728–33
- The Victorian Infant Collaborative Study Group. Surgery and the tiny baby: Sensorineural outcome at 5 years of age. *J Paediatr Child Health* 1996; 32:167–72
- Kalkman CJ, Peelen L, Moons KG, Veenhuizen M, Bruens M, Sinnema G, de Jong TP: Behavior and development in children and age at the time of first anesthetic exposure. *ANESTHESIOLOGY* 2009; 110:805–12
- Wilder RT, Flick RP, Sprung J, Katusic SK, Barbaresi WJ, Mickelson C, Gleich SJ, Schroeder DR, Weaver AL, Warner DO: Early exposure to anesthesia and learning disabilities in a population-based birth cohort. *ANESTHESIOLOGY* 2009; 110:796–804
- Flick RP, Katusic SK, Colligan RC, Wilder RT, Voigt RG, Olson MD, Sprung J, Weaver AL, Schroeder DR, Warner DO: Cognitive and behavioral outcomes after early exposure to anesthesia and surgery. *Pediatrics* 2011; 128:e1053–61
- Bartels M, Althoff RR, Boomsma DI: Anesthesia and cognitive performance in children: No evidence for a causal relationship. *Twin Res Hum Genet* 2009; 12:246–53
- Hansen TG, Pedersen JK, Henneberg SW, Pedersen DA, Murray JC, Morton NS, Christensen K: Academic performance in adolescence after inguinal hernia repair in infancy: A nationwide cohort study. *ANESTHESIOLOGY* 2011; 114:1076–85
- DiMaggio C, Sun LS, Kakavouli A, Byrne MW, Li G: A retrospective cohort study of the association of anesthesia and hernia repair surgery with behavioral and developmental disorders in young children. *J Neurosurg Anesthesiol* 2009; 21:286–91
- DiMaggio C, Sun LS, Li G: Early childhood exposure to anesthesia and risk of developmental and behavioral disorders in a sibling birth cohort. *Anesth Analg* 2011; 113:1143–51
- Jevtovic-Todorovic V: General anesthetics and the developing brain: Friends or foes? *J Neurosurg Anesthesiol* 2005; 17:204–6
- Clancy B, Darlington RB, Finlay BL: Translating developmental time across mammalian species. *Neuroscience* 2001; 105:7–17
- Clancy B, Finlay BL, Darlington RB, Anand KJ: Extrapolating brain development from experimental species to humans. *Neurotoxicology* 2007; 28:931–7
- Hoover HD, Dunbar SB, Frisbie DA, Oberley KR, Ordman VL, Naylor RJ, Bray GB, Lewis JC, Qualls AL, Mengeling MA, Shannon GP: The Iowa Tests: Guide to Research and Development: Forms A and B: Levels 5–14. Itasca: Riverside Publishing; 2003
- Gurka MJ, LoCasale-Crouch J, Blackman JA: Long-term cognition, achievement, socioemotional, and behavioral development of healthy late-preterm infants. *Arch Pediatr Adolesc Med* 2010; 164:525–32
- Slewa-Younan S, Baguley IJ, Herisceanu R, Cameron ID, Pitsovas V, Mudaliar Y, Nayyar V: Do men and women differ in their course following traumatic brain injury? A preliminary prospective investigation of early outcome. *Brain Inj* 2008; 22:183–91
- Hamilton BE, Martin JA, Ventura SJ: Births: Preliminary data for 2006. *Natl Vital Stat Rep* 2007; 56:1–18
- Flick RP, Wilder RT, Sprung J, Katusic SK, Voigt R, Colligan R, Schroeder DR, Weaver AL, Warner DO: Anesthesia and cognitive performance in children: No evidence for a causal relationship. Are the conclusions justified by the data? Response to Bartels *et al.*, 2009. *Twin Res Hum Genet* 2009; 12:611–2
- Wypij D, Newburger JW, Rappaport LA, duPlessis AJ, Jonas RA, Wernovsky G, Lin M, Bellinger DC: The effect of duration of deep hypothermic circulatory arrest in infant heart surgery on late neurodevelopment: The Boston Circulatory Arrest Trial. *J Thorac Cardiovasc Surg* 2003; 126:1397–403
- Bellinger DC, Wypij D, duPlessis AJ, Rappaport LA, Jonas RA, Wernovsky G, Newburger JW: Neurodevelopmental status at eight years in children with dextro-transposition of the great arteries: The Boston Circulatory Arrest Trial. *J Thorac Cardiovasc Surg* 2003; 126:1385–96
- Mahle WT, Clancy RR, Moss EM, Gerdes M, Jobes DR, Wernovsky G: Neurodevelopmental outcome and lifestyle assessment in school-aged and adolescent children with hypoplastic left heart syndrome. *Pediatrics* 2000; 105:1082–9
- Hövels-Gurich HH, Konrad K, Skorzewski D, Nacken C,

- Minkenberg R, Messmer BJ, Seghaye MC: Long-term neurodevelopmental outcome and exercise capacity after corrective surgery for tetralogy of Fallot or ventricular septal defect in infancy. *Ann Thorac Surg* 2006; 81:958–66
26. Hövels-Gurich HH, Seghaye MC, Däbritz S, Messmer BJ, von Bernuth G: Cognitive and motor development in preschool and school-aged children after neonatal arterial switch operation. *J Thorac Cardiovasc Surg* 1997; 114:578–85
  27. Hövels-Gurich HH, Seghaye MC, Schnitker R, Wiesner M, Huber W, Minkenberg R, Kotlarek F, Messmer BJ, von Bernuth G: Long-term neurodevelopmental outcomes in school-aged children after neonatal arterial switch operation. *J Thorac Cardiovasc Surg* 2002; 124:448–58
  28. Dessens AB, Cohen-Kettenis PT, Mellenbergh GJ, Koppe JG, van De Poll NE, Boer K: Association of prenatal phenobarbital and phenytoin exposure with small head size at birth and with learning problems. *Acta Paediatr* 2000; 89:533–41
  29. Ludman L, Spitz L, Wade A: Educational attainments in early adolescence of infants who required major neonatal surgery. *J Pediatr Surg* 2001; 36:858–62

## ANESTHESIOLOGY REFLECTIONS FROM THE WOOD LIBRARY-MUSEUM

### The Plaque to “Ethylene-Oxygen Anesthesia”



As pictured in the close-up (*left*), a commemorative plaque was placed in the operating room in Chicago's Presbyterian Hospital in which on "March 14, 1923 ethylene-oxygen anesthesia was administered for the first time to a human surgical patient." Witnessed by at least 10 physicians, that anesthetic was administered by Dr. Lester R. Dragstedt (*not pictured*). The photographed woman (*right*) was Dr. Mary Lyons, who, along with Dr. Isabella Herb, would become famous as physician-anesthetists who pioneered anesthetic use of ethylene. Next to Dr. Lyons is physiologist Arno Luckhardt, M.D., who must have been relieved that — just 3 days before the American Medical Association published his article on self-experimenting with ethylene — his University of Chicago colleagues had been the first to successfully anesthetize a human surgical patient with that gas. (Copyright © the American Society of Anesthesiologists, Inc.)

George S. Bause, M.D., M.P.H., Honorary Curator, ASA's Wood Library-Museum of Anesthesiology, Park Ridge, Illinois, and Clinical Associate Professor, Case Western Reserve University, Cleveland, Ohio. UJYC@aol.com.