ORIGINAL CLINICAL RESEARCH REPORT

Prenatal Exposure to General Anesthesia and Childhood Behavioral Deficit

Caleb Ing, MD, MS,* Ruth Landau, MD,† David DeStephano, MPH,† Caleb H. Miles, PhD,‡ Britta S. von Ungern-Sternberg, MD, PhD,§‖¶ Guohua Li, MD, DrPH,* and Andrew J. O. Whitehouse, PhD#

BACKGROUND: Exposure to surgery and anesthesia in early childhood has been found to be associated with an increased risk of behavioral deficits. While the US Food and Drug Administration (FDA) has warned against prenatal exposure to anesthetic drugs, little clinical evidence exists to support this recommendation. This study evaluates the association between prenatal exposure to general anesthesia due to maternal procedures during pregnancy and neuropsychological and behavioral outcome scores at age 10.

METHODS: This is an observational cohort study of children born in Perth, Western Australia, with 2 generations of participants contributing data to the Raine Study. In the Raine Study, the first generation (Gen1) are mothers enrolled during pregnancy, and the second generation (Gen2) are the children born to these mothers from 1989 to 1992 with neuropsychological and behavioral tests at age 10 (n=2024). In the primary analysis, 6 neuropsychological and behavioral tests were evaluated at age 10: Raven's Colored Progressive Matrices (CPM), McCarron Assessment of Neuromuscular Development (MAND), Peabody Picture Vocabulary Test (PPVT), Symbol Digit Modality Test (SDMT) with written and oral scores, Clinical Evaluation of Language Fundamentals (CELF) with Expressive, Receptive, and Total language scores, and Child Behavior Checklist (CBCL) with Internalizing, Externalizing, and Total behavior scores. Outcome scores of children prenatally exposed to general anesthesia were compared to children without prenatal exposure using multivariable linear regression models adjusting for demographic and clinical covariates (sex, race, income, and maternal education, alcohol or tobacco use, and clinical diagnoses: diabetes, epilepsy, hypertension, psychiatric disorders, or thyroid dysfunction). Bonferroni adjustment was used for the 6 independent tests in the primary analysis, so a corrected P value <.0083 (P = .05 divided by 6 tests, or a 99.17% confidence interval [CI]) was required for statistical significance.

RESULTS: Among 2024 children with available outcome scores, 22 (1.1%) were prenatally exposed to general anesthesia. Prenatally exposed children had higher CBCL Externalizing behavioral scores (score difference of 6.1 [99.17% CI, 0.2-12.0]; P = .006) than unexposed children. Of 6 tests including 11 scores and subscores, only CBCL Externalizing behavioral scores remained significant after multiple comparisons adjustment with no significant differences found in any other score.

CONCLUSIONS: Prenatal exposure to general anesthetics is associated with increased externalizing behavioral problems in childhood. However, given the limitations of this study and that avoiding necessary surgery during pregnancy can have significant detrimental effects on the mother and the child, further studies are needed before changes to clinical practice are made. (Anesth Analg XXX;XXX:00–00)

KEY POINTS

- **Question:** Is prenatal exposure to surgery and anesthesia associated with long-term neurodevelopmental deficits?
- **Findings:** Children with prenatal exposure to general anesthesia due to maternal surgery during pregnancy had more externalizing behavioral problems at 10 years of age compared to unexposed children.
- **Meaning:** This provides preliminary epidemiologic evidence that anesthetic exposure during pregnancy may be associated with a child's long-term development.

From the *Department of Anesthesiology and Epidemiology, Columbia University College of Physicians and Surgeons and Mailman School of Public Health, New York, New York; †Department of Anesthesiology, Columbia University College of Physicians and Surgeons, New York, New York; ‡Department of Biostatistics, Mailman School of Public Health, New York, New York; §Division of Emergency Medicine, Anaesthesia and Pain Medicine, Medical School, The University of Western Australia, Perth, Australia; ‖Department of Anaesthesia and Pain Management, Perth Children's Hospital, Perth, Australia; ¶Team Perioperative Medicine, Telethon Kids Institute, Perth, Australia; and #Telethon Kids Institute, University of Western Australia, Perth, Australia.

Accepted for publication December 9, 2020.

DOI: 10.1213/ANE.0000000000005389

Funding: The Raine Study is funded by project and program grants from the National Health and Medical Research Council of Australia (NHMRC) (Canberra, Australia). Core management funding is provided by the Raine Medical Research Foundation, the Telethon Kids Institute, the University of Western Australia (UWA), the Women and Infants Research Foundation, Curtin University, Murdoch University, Edith Cowan University, and the University of Notre Dame Australia. C.I. is supported by a Herbert Irving Scholars Award as well as the Agency for Healthcare Research and Quality (AHRQ) under award number R01HS026493. The content is solely the responsibility of the authors and does not necessarily represent the official views of the AHRQ. A.J.O.W. is supported by an Investigator Grant from the National Health and Medical Research Council (APP1173896). B.S.v.U.-S. is partly supported by the Perth Children's Hospital Foundation, the Stan Perron Charitable Trust (Australia), and the Callahan Estate (Australia).

GLOSSARY

aOR = adjusted odds ratios; **CBCL** = Child Behavior Checklist; **CELF** = Clinical Evaluation of Language Fundamentals; **CI** = confidence interval; **CPM** = Raven's Colored Progressive Matrices; **FDA** = US Food and Drug Administration; **GA** = gestational age; **Gen1** = first generation in the Raine Study; **Gen2** = second generation in the Raine Study; **IV** = intravenous; **MAND** = McCarron Assessment of Neuromuscular Development; **MASK** = Mayo Anesthesia Safety in Kids; **NDI** = Neurodevelopmental Index; **PANDA** = Pediatric Anesthesia NeuroDevelopment Assessment; **PPVT** = Peabody Picture Vocabulary Test; **SDMT** = Symbol Digit Modality Test

Millions of children receive anesthesia each year for surgical and diagnostic procedures. Questions have emerged about the safety of anesthetics in children given that exposure to anesthetic agents during brain maturation disrupts neurodevelopment in preclinical models,[1–3] and early childhood exposure to anesthesia may be associated with neurocognitive deficits later in life.[4–6] Based on these concerns, the US Food and Drug Administration (FDA) issued a warning in 2016 about the neurodevelopmental effects of anesthetic drugs in "children younger than 3 years or in pregnant women during their third trimester."[7] However, long-term neurodevelopmental outcomes following prenatal exposures have yet to be evaluated in a clinical study, and the recommendation from the FDA is based in large part on animal models. The scientific premise, though, is plausible as all general anesthetic drugs have been found to cross the placenta,[8] and the prenatal period is characterized by high fetal sensitivity to neurotoxic agents throughout all trimesters of gestation.[9] In addition, the prenatal neurodevelopmental time period has specifically been associated with peak brain vulnerability in preclinical studies of anesthetic neurotoxicity.[3,10] Therefore, the primary purpose of this study is to evaluate whether prenatal exposure to general anesthetic agents due to maternal need for surgery and anesthesia during pregnancy is associated with differences in neuropsychological or behavioral survey scores at age 10. Additional secondary analyses were performed to explore the clinical significance of any score differences, as well as the contribution of perinatal and postnatal factors that may occur after prenatal exposure.

METHODS

This study was approved by the Columbia University Institutional Review Board, with the requirement of written informed consent waived. Data collection and storage at each age were approved by Ethics Committees at King Edward Memorial Hospital, Princess Margaret Hospital, and the University of Western Australia.

The Raine Study Cohort

The Raine Study is an established birth cohort in Perth, Western Australia, with 2 generations of participants contributing data: the first generation (Gen1), who are the mothers enrolled during pregnancy, and the second generation (Gen2), consisting of their 2868 children born from 1989 to 1992.[11] As part of the study, detailed demographic and medical data were collected prenatally and at birth from medical records and parental self-report. Children were assessed for the presence of medical illnesses at various ages after birth and had comprehensive neuropsychological testing performed at the age 10 follow-up.[5]

Prenatal Exposure

Exposure was defined as maternal exposure to general anesthesia during pregnancy and was determined based on questionnaires completed by the mothers during 2 antenatal study visits at approximately 18 and 34 weeks of gestational age (GA). In the survey at 18 weeks of GA, mothers were specifically asked, "Since you have been pregnant, have you had a general anesthetic?" The mothers were asked to describe the operation, as well as either the approximate date of the procedure or the GA of the child at the time. The same questions were asked on the 34-week survey, which specifically asked about general anesthesia in the time since the 18-week visit (Supplemental Digital Contents 1–3, Supplemental Online Content and Supplemental Figures 1 and 2, http://links.lww.com/AA/D352, http://links.lww.com/AA/D353, http://links.lww.com/AA/D354). For the present study, the written survey responses were reviewed and abstracted from the original study questionnaires. Since there may be confusion on the part of the mothers regarding what may constitute a general anesthetic, the responses by the mothers were independently reviewed. Consensus on the likelihood of the mother receiving a general anesthetic based on the type and timing of the procedure during gestation was reached in discussions between an obstetric anesthesiologist and 2 pediatric anesthesiologists, one of whom has first-hand knowledge of clinical anesthetic practices

The authors declare no conflicts of interest.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (www.anesthesia-analgesia.org).

Reprints will not be available from the authors.

Address correspondence to Caleb Ing, MD, MS, Department of Anesthesiology and Epidemiology, Columbia University College of Physicians and Surgeons and Mailman School of Public Health, 622 W 168th St BHN 4-440, New York, NY 10032. Address e-mail to ci2119@cumc.columbia.edu.

in Western Australia (R.L., C.I., and B.S.v.U.-S.). For any operations where the likely type of anesthetic was unclear, advice was sought from additional obstetric anesthesiologists who administered clinical care in Western Australia during the time period when the women were enrolled. Procedures were classified as (1) likely general anesthetic, (2) likely intravenous (IV) sedation, or (3) unlikely to have required a general anesthetic or IV sedation. For the primary analysis, all maternal data forms that lacked information on the type of procedure were classified as "Unlikely to have required general anesthesia or IV sedation."

Neurodevelopmental Outcomes

A range of cognitive domains was evaluated using 6 directly assessed neuropsychological tests and parental surveys. Cognition was assessed by the Symbol Digit Modality Test (SDMT) and the Raven's Colored Progressive Matrices (CPM) with the SDMT assessing visual tracking, attention, and fine motor skills and generating written and oral scores, and the CPM measuring global cognitive performance, nonverbal intelligence, and visuospatial functions.[12,13] Motor function was assessed using the McCarron Assessment of Neuromuscular Development (MAND), which evaluated both fine and gross motor tasks and generated a Neurodevelopmental Index Score.[14] The Clinical Evaluation of Language Fundamentals (CELF) is a language test that assesses higher-order semantic, grammatical, and verbal memory abilities and generates 3 scores. The CELF-R is the receptive language score measuring listening comprehension, CELF-E is the expressive language score tracking speaking ability, and the CELF-T represents total language ability.[15] The Peabody Picture Vocabulary Test (PPVT) is a receptive listening vocabulary test that also assesses language.[16] Behavioral problems were measured by the Child Behavior Checklist (CBCL), a 118-item informant-report questionnaire evaluating internalizing problems such as depression and somatic complaints, as well as externalizing problems such as aggressive behavior and rule-breaking. In addition to internalizing and externalizing scores, the CBCL also generates a total behavior score.[17] As opposed to the other neuropsychological tests in which higher scores show better performance, for CBCL, higher scores show more behavioral problems, with scores >60 considered as clinical deficit.[18,19] The CBCL tests were completed by the child's caregiver, and since they did not require the presence of the child for testing, they were completed at a higher rate than other tests.

Statistical Analysis

Children of mothers classified as being likely exposed to general anesthesia during pregnancy were classified as exposed and compared to children of mothers who reported no history of prenatal exposure to anesthesia. Crude mean scores for exposed and unexposed children were initially evaluated. As the primary analysis, multivariable linear regression was used to evaluate score differences after adjusting for demographic and clinical covariates (sex, race, maternal school level, income, alcohol or tobacco use during pregnancy, as well as the presence of any maternal clinical diagnoses: diabetes, epilepsy, hypertension, psychiatric disorders, or thyroid dysfunction). A Bonferroni adjustment was used for multiple comparisons. The Bonferroni adjustment is known to be conservative, particularly when outcomes are correlated, which may jeopardize sensitivity to detect true signals.[20] Given the high degree of correlation between scores and subscores for specific neuropsychological tests, in this study, we considered there to be 6 independent tests, 1 for each of the 6 individual neuropsychological and behavioral tests evaluated. For this adjustment, the significance level (α) was .05 for the 6 independent tests. Therefore, the corrected P value threshold for significance was set to .05/6 or $P < .0083$ and a confidence interval (CI) of 99.17% was used for our primary analysis. This corrected P value threshold of $P < .0083$ was applied to the secondary analysis evaluating risk of crossing a clinical threshold. The mediation analysis was an exploratory analysis and a P value threshold of $P < .05$ was used.

Evaluation of Risk of Crossing a Clinical Threshold

Two secondary analyses were performed to aid in the interpretability of the primary analysis. The first explores the clinical significance of any score differences found in the primary analysis by evaluating the increased risk of any child crossing an established clinical threshold for deficit. For CBCL, a frequently used clinical cutoff, including other studies evaluating children after exposure to anesthesia, is a score >60, which corresponds to a score that is 1 standard deviation worse than the normalized mean score in the population used to develop the instrument.[18,19] For the other assessments where clinical deficit has been defined at various levels, to maintain consistency, scores worse than 1 standard deviation from the mean of the present population were defined as clinical deficit. In this analysis, multivariable logistic regression models were used to calculate adjusted odds ratios (aOR) of deficit.

Additional Descriptive Evaluations

To evaluate whether children with prenatal exposure to general anesthesia had higher rates of comorbidity, health care utilization using a previously defined resource utilization variable in children with prenatal

anesthetic exposure was reported.[21] To evaluate neuropsychological and behavioral outcome scores based on the trimester of exposure, crude trimester-specific scores were also reported based on the reported date or GA of exposure. Mothers who did not report specific dates of exposure were excluded from this analysis.

Sensitivity Analyses

Missing covariates were initially coded as separate levels to maximize the total number of children included in the analysis, but as sensitivity analyses, complete case analyses were also performed. In our primary analysis, determination of general anesthetic exposure was based on the expert opinion of 3 reviewers. Since these reviewers did not have access to the anesthetic records, the determinations were made based on procedure type and GA of the child at the time of the procedure, and therefore, there may be a risk for misclassification. This possibility was explored in an additional sensitivity analysis by repeating the primary multivariable linear regression analysis but changing the prenatal exposure variable to include any mother who reported receiving a general anesthetic regardless of the documented procedure. In this analysis, the threshold for significance was maintained at $P < .0083$.

Analysis of Potential Mediators

Surgery and anesthesia during pregnancy are associated with perinatal factors and complications and postnatal events, which may affect long-term neurodevelopmental outcomes.[22] If these factors occur after exposure and potentially as a result of the exposure, adjusting for these postexposure variables may result in overadjustment bias.[23] Therefore, to evaluate the association of these factors on outcomes that are significantly associated with prenatal exposure to general anesthesia, a mediation analysis was performed as a secondary exploratory analysis with perinatal and postnatal factors evaluated as potential mediators between prenatal general anesthetic exposure and neurodevelopmental outcomes. In the outcomes showing a significant difference, the "total effect" of the prenatal exposure is the effect of the prenatal exposure on the outcome while ignoring the mediator variables.[24] This is the effect we estimated in the primary analysis and is composed of the "direct effect," which is the effect of the exposure on the outcome that is not transmitted through the mediator, and the "indirect effect," which is the effect that goes through the mediator. For outcomes not showing a statistically significant total effect, the importance of the indirect effect of a mediator is limited and therefore these outcomes were not evaluated. The potential mediators evaluated included factors in the child assessed at the time of birth such as prematurity, low birth weight, intrauterine growth retardation, maternal factors at the time of delivery such as need for cesarean delivery or epidural anesthesia use, as well as postnatal surgery before age 3 years old. Due to the presence of sparse data in some covariates, the mediation analysis was performed in children with complete covariate data (complete case analysis model). Despite this, sparse data were still found in the models evaluating prematurity as a mediator, and as a result, in those models, the maternal heart disease covariate was removed. We calculated "percent mediated" by dividing the mediated effect by the total effect of the exposure on the outcome. Since mediation analysis may be susceptible to bias in the presence of exposure-mediator interaction when this interaction is not included in the regression model, interactions between the mediator and exposure variables were included in all models.[25] We report the "percent due to interaction," which divides the average portion of the total effect that is attributable to interaction by the total effect.[26] For additional details and assumptions made in the mediation analysis, please see (Supplemental Digital Content 4, Methods, http://links.lww.com/AA/D355). SAS 9.4 was used for all analyses with the mediation analysis performed using the CAUSALMED procedure.

RESULTS

The Raine Study cohort consists of 2868 live births in Western Australia from 1989 to 1992 with mothers enrolled during pregnancy. Of these children, 45 had mothers who completed a survey at 18 or 34 weeks stating that they had received a general anesthetic during pregnancy. On review of the maternal procedures recorded on the survey forms, based on the likely anesthetic used for the maternal surgery, 30 children were judged to likely have had a prenatal exposure to general anesthesia, 1 was likely to have had a prenatal exposure to IV sedation, and 14 were unlikely to have had a prenatal exposure to general anesthesia or IV sedation (Figure 1). The surgical procedures performed on the expecting mothers were mostly minor in nature and took place at GAs ranging from 3 to 34 weeks (Table 1). Most mothers reported either the date of surgery or the GA of the child. Some mothers, however, did not report a date and the only information regarding the timing of the operation was whether the survey form was filled during the 18- or 34-week follow-up visit.

There were a total of 2024 children with some neuropsychological testing available at age 10 years, of which 22 (1.1%) were likely exposed prenatally to general anesthesia, 1 (0.05%) was likely exposed prenatally to IV sedation, 7 (0.3%) were unlikely to have been exposed to general anesthesia or IV sedation despite a reported exposure, and 1994 (98.5%)

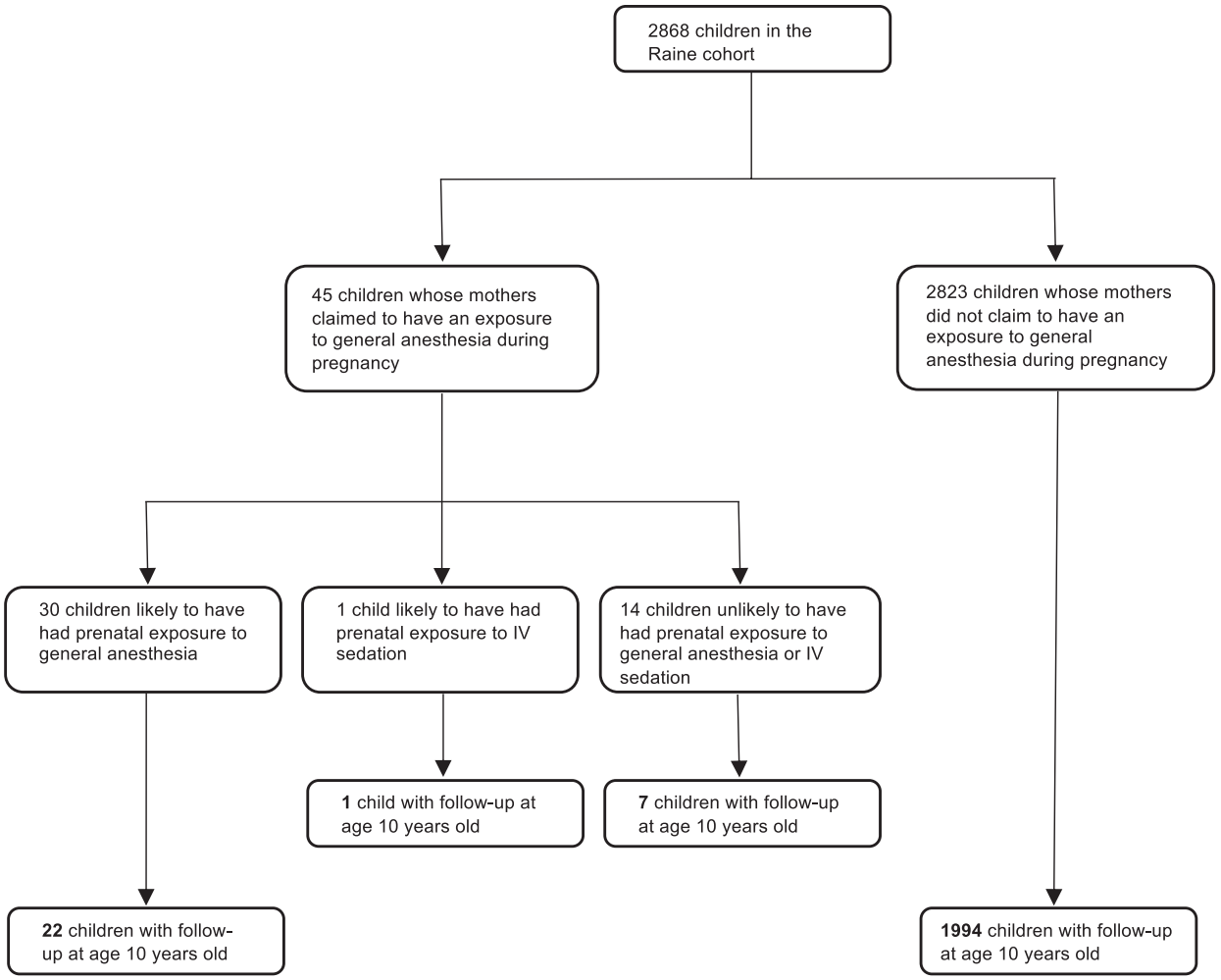

Figure 1. Participant flow diagram. IV indicates intravenous.

had no reported prenatal exposure to general anesthesia. Mothers of children with prenatal exposure to general anesthesia were less likely to have a college or university degree and more likely to smoke cigarettes and use alcohol during pregnancy than mothers of children with no prenatal exposure to anesthetics (Table 2).

Neuropsychological Test Score Differences Between Prenatally Exposed and Unexposed Children

Crude mean differences in test scores between prenatally exposed and unexposed children were evaluated (Supplemental Digital Content 5, Table 1, http://links.lww.com/AA/D356). After adjusting for demographic and clinical covariates using multivariable linear regression, compared to unexposed children, those with prenatal exposure to general anesthesia had significantly higher/worse CBCL Externalizing behavioral scores (score difference of 6.1 [99.17% CI, 0.2-12.0]; $P = .006$; Figure 2). No significant differences were found in any other score.

Clinical Threshold Analysis

To evaluate the clinical implications of these score differences, all scores were evaluated to see if they crossed a threshold for clinical deficit. Of the prenatally exposed children, 36.4% crossed a threshold for clinical deficit in CBCL Externalizing behavioral scores compared to 10.5% of children with no history of prenatal exposure (Table 3). After adjusting for demographic and clinical covariates using multivariable logistic regression, prenatal anesthetic exposure was associated with an increased risk of clinical deficit in externalizing behavioral scores (aOR = 4.6 [99.17% CI, 1.3-16.0]; $P = .001$). No other outcome scores showed statistically significant differences following multiple comparisons adjustment. For some outcomes, no children with prenatal exposure to anesthesia were in the deficit range, so those outcomes could not be evaluated.

Additional Descriptive Evaluations

When evaluating comorbidity in these children after birth, children with prenatal exposure to general anesthesia had similar rates of health care resource

Table 1. Procedures Performed in the 46 Mothers Reporting General Anesthetic Exposure During Pregnancy

	Procedure	Gestational age[a] (wk)	Trimester
Likely general anesthetic	Abscess lanced	10	1
	Abscess resection	7	1
	Appendectomy	9	1
	Appendectomy	11	1
	Appendectomy	23	2
	Appendectomy	17	2
	Appendectomy	34-wk survey	-
	Appendectomy	3	1
	Bartholin's abscess	11	1
	Bartholin's cyst removal	29 and 34	3
	Bartholin's cyst removal	34-wk survey	-
	Bone graft	18-wk survey	-
	Carpal tunnel surgery	3	1
	Cervical suture	13	2
	Cervical suture	18	2
	Cervical suture	18	2
	Cervical suture	14.5	2
	Cervical suture	16	2
	Drain vaginal abscess	7	1
	Gastroscopy and sigmoidoscopy	4	1
	Hemorrhoid surgery	28	3
	Laparoscopy	3	1
	Laparoscopy	3	1
	Lymph node resection	4.5	1
	Ovarian cyst removal	7	1
	Ovarian cyst removal	18	2
	Pins placed in wrist	29	3
	Tonsillectomy	"Was just pregnant"	1
	Tooth extraction	36	3
	Wisdom tooth extraction	5	1
Likely IV sedation	Pleura effusions drained	34-wk survey	-
Unlikely to have required general anesthetic or IV sedation	Amniocentesis	18	2
	Dental	18-wk survey	-
	Dental	18-wk survey	-
	Dental	11.5	1
	Dental	34-wk survey	-
	Dermabrasion	3	1
	Mole resection	34-wk survey	-
	Not stated	18-wk survey	-
	Not stated	18-wk survey	-
	Not stated	18-wk survey	-
	Not stated	18-wk survey	-
	Not stated	19	2
	Not stated	34-wk survey	-
	Suture on foot	18-wk survey	-

Abbreviation: IV, intravenous.
[a]Reported the prenatal survey where the operation was disclosed when procedure date and gestational age were unavailable.

Table 2. Characteristics of the Mothers and Children

	Likely exposed to general anesthetic n (%)	No report of exposure to general anesthetic n (%)
Sex		
Female	9 (40.9)	961 (48.2)
Male	13 (59.1)	1033 (51.8)
Race		
White	21 (95.5)	1767 (88.6)
Other	1 (4.5)	186 (9.3)
Missing	0 (0)	41 (2.1)
Maternal education		
None	15 (68.2)	918 (46)
Trade certificate professional registration or other	4 (18.2)	469 (23.5)
College or university degree	3 (13.6)	566 (28.4)
Missing	0 (0)	41 (2.1)
Income		
<$7000	2 (9.1)	126 (6.3)
$7000–$23,999	7 (31.8)	578 (29)
$24,000–$35,999	7 (31.8)	483 (24.2)
≥$36,000	6 (27.3)	681 (34.2)
Missing	0 (0)	126 (6.3)
Smoking during pregnancy		
None	10 (45.5)	1450 (72.7)
1–10 cigarettes daily	6 (27.3)	197 (9.9)
>10 cigarettes daily	3 (13.6)	179 (9)
Missing	3 (13.6)	168 (8.4)
Alcohol use during pregnancy		
Never	10 (45.5)	1073 (53.8)
Once a week or less	9 (40.9)	631 (31.6)
More than once a week	0 (0)	105 (5.3)
Missing	3 (13.6)	185 (9.3)
Maternal diabetes		
No	22 (100)	1914 (96)
Yes	0 (0)	39 (2)
Missing	0 (0)	41 (2.1)
Maternal epilepsy		
No	22 (100)	1917 (96.1)
Yes	0 (0)	36 (1.8)
Missing	0 (0)	41 (2.1)
Maternal heart disease		
No	22 (100)	1910 (95.8)
Yes	0 (0)	42 (2.1)
Missing	0 (0)	42 (2.1)
Maternal psychiatric disorder		
No	21 (95.5)	1910 (95.8)
Yes	1 (4.5)	43 (2.2)
Missing	0 (0)	41 (2.1)
Maternal thyroid dysfunction		
No	21 (95.5)	1905 (95.5)
Yes	1 (4.5)	48 (2.4)
Missing	0 (0)	41 (2.1)

utilization in the first 10 years of life when compared to children with no exposure to anesthetics (Supplemental Digital Content 6, Table 2, http://links.lww.com/AA/D357). In evaluating trimester-specific exposures, the majority of exposures occurred during the first trimester with 10 or 13 children exposed in the first trimester and evaluated at age 10. For children exposed in the second trimester, 5–6 were evaluated at age 10, and for children exposed in the third trimester, only 2 were evaluated at age 10. The highest mean CBCL scores were seen in the children exposed in the first trimester. However, given the limited sample size, no statistical tests were performed, and these results should be interpreted with caution (Supplemental Digital Content 7, Table 3, http://links.lww.com/AA/D358).

Sensitivity Analyses

When performing a complete case analysis, prenatally exposed children still had higher/worse CBCL

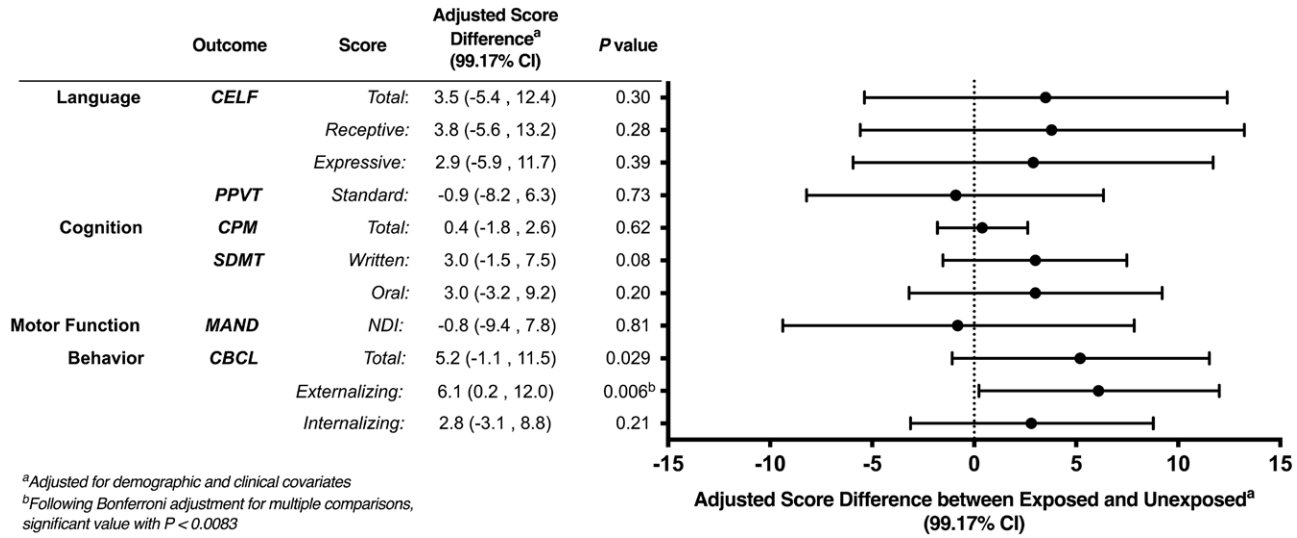

Figure 2. Neuropsychological score differences between children with and without prenatal exposure to general anesthesia. The scores included the CELF, PPVT, Raven's CPM, SDMT, MAND NDI, and CBCL. Score differences are adjusted for demographic and clinical covariates (sex, race, maternal school level, income, alcohol or tobacco use during pregnancy, and presence of any maternal clinical diagnoses: diabetes, epilepsy, heart disease, psychiatric disorders, or thyroid dysfunction). CBCL indicates Child Behavior Checklist; CELF, Clinical Evaluation of Language Fundamentals; CI, confidence interval; CPM, Colored Progressive Matrices; MAND, McCarron Assessment of Neuromuscular Development; NDI, Neurodevelopmental Index; PPVT, Peabody Picture Vocabulary Test; SDMT, Symbol Digit Modality Test.

Table 3. Odds Ratios of Clinical Deficit Associated With Prenatal Anesthetic Exposure

			Likely exposed to general anesthetic during pregnancy			No report of exposure to general anesthetic during pregnancy				
Neuropsychological domain	Outcome	Score	Total (n)	Deficit (n)	Deficit (%)	Total (n)	Deficit (n)	Deficit (%)	Adjusted OR[a] (99.17% CI)	*P* value[a]
Language	CELF	Total	19	0	0	1600	242	15.1	-	-
		Receptive	19	0	0	1600	220	13.8	-	-
		Expressive	19	2	10.5	1600	254	15.9	0.5 (0.1-3.5)	.31
	PPVT	Standard	18	3	16.7	1482	234	15.8	0.9 (0.2-5.3)	.93
Cognition	CPM	Total	19	3	15.8	1601	256	16	0.8 (0.1-4.2)	.67
	SDMT	Written	19	0	0	1591	209	13.1	-	-
		Oral	19	2	10.5	1591	249	15.7	0.6 (0.1-4.2)	.45
Motor function	MAND	NDI	18	7	38.9	1584	266	16.8	3.2 (0.9-12.2)	.019
Behavior	CBCL	Total	22	8	36.4	1967	264	13.4	3.4 (1.0-11.6)	.009
		Externalizing	22	8	36.4	1967	206	10.5	4.6 (1.3-16.0)	.001[b]
		Internalizing	22	7	31.8	1967	283	14.4	2.5 (0.7-8.8)	.05

Abbreviations: CBCL, Child Behavior Checklist; CELF, Clinical Evaluation of Language Fundamentals; CI, confidence interval; CPM, Colored Progressive Matrices; MAND, McCarron Assessment of Neuromuscular Development; OR, odds ratio; PPVT, Peabody Picture Vocabulary Test; SDMT, Symbol Digit Modality Test.

[a]Adjusted for demographic and clinical covariates (sex, race, maternal school level, income, alcohol or tobacco use during pregnancy, and presence of any maternal clinical diagnoses: diabetes, epilepsy, heart disease, psychiatric disorders, or thyroid dysfunction).

[b]Statistically significant at $P < .0083$.

Externalizing behavioral scores (score difference of 6.7 [99.17% CI, 0.5-13.0]; $P = .004$) than unexposed children. Like the primary analysis, no statistically significant differences were found in any other score after multiple comparisons adjustment. As a sensitivity analysis, we redefined prenatal exposure as any mother claiming that they had a general anesthetic during pregnancy regardless of their procedure type, which included mothers having dental and minor skin procedures that are unlikely to require general anesthesia. In this analysis, compared to unexposed children, exposed children were found to have a CBCL Externalizing score difference of 5.0 (99.17% CI, −0.1 to 10.2; $P = .010$). After using this broader definition of prenatal exposure, the absolute adjusted mean score difference for CBCL Externalizing scores was not as large as that found in the primary analysis and was not statistically significant after multiple comparisons adjustment (Supplemental Digital Content 8, Table 4, http://links.lww.com/AA/D359). No other outcome scores were statistically significant after multiple comparisons adjustment.

Evaluation of Potential Mediators

When evaluating the 6 potential mediators, in the children with prenatal exposure to general anesthesia, a higher rate of cesarean delivery (31.8% vs 20.9%) and epidural anesthetic use (63.6% vs 47.7%)

was seen compared to children with no history of prenatal exposure (Supplemental Digital Content 9, Table 5, http://links.lww.com/AA/D360). The other variables were similar between the children with and without prenatal exposure. The percentage of the total effect mediated by each of these potential mediators was evaluated and none were found to be statistically significant. However, when evaluating the percentage of the total effect due to interaction, the interaction between cesarean delivery and exposure was found to be statistically significant (percent due to interaction: 73.1; 95% CI, 4.8-141.3; $P = .036$; Table 4).

DISCUSSION

Based on our results, prenatal exposure to general anesthesia due to maternal surgery during pregnancy was associated with worse childhood externalizing behavioral problem scores and a 4.6 times increased risk of crossing a threshold for a clinically significant deficit in externalizing behavioral problems. However, no differences were found in other neurocognitive domains. That the differences were found specifically in behavioral scores is consistent with other studies evaluating neuropsychological and behavioral outcomes in children after postnatal anesthetic exposures. In the Pediatric Anesthesia NeuroDevelopment Assessment (PANDA) study, the only outcomes showing significant differences were the CBCL Total and Internalizing scores, which were worse in anesthetic-exposed children when compared to their unexposed siblings.[18] The Mayo Anesthesia Safety in Kids (MASK) study also identified significantly worse CBCL Total, Internalizing, and Externalizing scores, but only after more than 1 anesthetic exposure.[19] The differences found in the PANDA study ranged from 2.7 to 3.2 points worse in the children with a single anesthetic exposure, while multiply exposed children in the MASK study had scores that were 2.9–4.8 points worse. The mean score differences found in the present study in the CBCL Total and Internalizing scores were 5.2 and 6.1 points worse (more than one-half of a standard deviation), which is at least equivalent with those found in children with postnatal multiple exposures.[19] The CBCL results from the PANDA and MASK studies, however, were not the primary outcomes in those studies and should therefore be interpreted with caution. The results from the present study, though, differ from those evaluating a cohort of children from the Raine Study with postnatal exposures. In a prior study, when evaluating children with postnatal exposures under the age of 3 years old, worse scores in language and abstract reasoning were found, while no significant differences in behavior were identified.[5]

A secondary analysis was performed evaluating outcomes after trimester-specific exposures. Given the

Table 4. Mediation Analyses of CBCL Externalizing Scores Using Complete Cases

	Mediators					
	Prematurity Estimate (95% CI)	Birth weight <2500 g Estimate (95% CI)	Intrauterine growth retardation Estimate (95% CI)	Cesarean delivery Estimate (95% CI)	Epidural anesthesia Estimate (95% CI)	Postnatal surgery Estimate (95% CI)
Total effect (score difference)	6.68 (1.54-11.82)	6.57 (1.53-11.62)	6.51 (−2.14 to 15.15)	6.88 (1.15-12.61)	6.67 (1.19-12.15)	6.99 (2.19-11.79)
Direct effect	6.66 (2.04-11.28)	6.63 (2.01-11.26)	6.45 (0.10-12.8)	4.86 (0.10-9.63)	4.86 (0.05-9.67)	7.06 (2.42-11.7)
Indirect effect	0.02 (−2.26 to 2.29)	−0.06 (−2.1 to 1.98)	0.06 (−2.97 to 3.08)	2.02 (−1.62 to 5.65)	1.81 (−1.43 to 5.06)	−0.07 (−1.35 to 1.21)
Percent mediated	0.2 (−33.7 to 34.2)	−0.9 (−32.2 to 30.4)	0.9 (−44.6 to 46.4)	29.3 (−14.8 to 73.4)	27.2 (−15.3 to 69.6)	−1.0 (−19.5 to 17.4)
Percent due to Interaction	16.7 (−16.3 to 49.6)	16.2 (−16.1 to 48.6)	2.3 (−114.9 to 119.5)	73.1 (4.8-141.3)	115.3 (−5.6 to 236.2)	−17.2 (−51.1 to 16.8)

Analyses of all mediators except for prematurity were adjusted for sex, race, maternal school level, income, alcohol or tobacco use during pregnancy, and presence of any maternal clinical diagnoses: diabetes, epilepsy, heart disease, psychiatric disorders, or thyroid dysfunction. The prematurity analysis was adjusted for all covariates except for maternal heart disease.
Abbreviations: CBCL, Child Behavior Checklist; CI, confidence interval.

small sample size, the conclusions that can be drawn from this are limited. However, it is worth noting that while the FDA only warns against third-trimester exposures,[7] in this cohort, the highest/worst mean CBCL scores were actually found in the children with a first-trimester exposure.

Cesarean deliveries and epidural use were present at a higher rate in the prenatally exposed children, but the results of the exploratory mediation analysis found that the CBCL Externalizing score differences in prenatally exposed children could not be significantly attributed to any of the mediators. However, there was significant interaction when evaluating cesarean deliveries, meaning that the association between CBCL Externalizing scores and prenatal exposure to anesthesia varied significantly based on whether the mothers received a cesarean delivery. Specifically, this association of worse scores with prenatal exposure was stronger among the children who had prenatal exposure and were delivered via cesarean delivery than those who had prenatal exposure and were not delivered via cesarean delivery. In evaluating the children with prenatal exposure and cesarean delivery, 5 of the 6 cesarean deliveries occurred electively due to prior cesarean deliveries in the mother or the positioning of the baby. In this exploratory mediation analysis, a clear reason for this interaction could not be identified.

This study has a number of limitations. Since anesthetic exposures during pregnancy in women are a relatively uncommon occurrence, it is difficult to assemble a large cohort of prenatally exposed children with long-term neuropsychological follow-up. Therefore, the first limitation is the small sample size. Despite the limited number of children assessed, significant differences in long-term behavioral function were still identified that are consistent with other studies of postnatal anesthetic exposure.[18,19,27] However, since statistical significance depends in large part on sample size, the lack of statistical significance on a given outcome does not preclude the presence of an effect in a larger cohort. Also, while the exploratory mediation analysis found an interaction between prenatal exposure and cesarean delivery, these results should be interpreted with caution given the limited number of children with both exposures and the lack of multiple comparisons adjustment in this analysis. A second limitation is the inability to distinguish between the effects of the anesthetic medication or uterine perfusion, inflammation, and any other stressor that the fetus may be exposed to during the maternal surgery. Given the lack of an anesthetic record, the types and doses of anesthetic medications, as well as any intraoperative complications, are unknown. While all procedures were relatively minor in nature, an additional consideration is that since these maternal exposures occurred between 1988 and 1992, the medications and monitoring techniques used in these mothers may differ from those commonly used in present-day practice. Despite the differences in clinical practice over time, these results are still relevant as the behavioral outcome signal found can be used to inform future studies. Third, since the likelihood of general anesthetic exposure was judged based on maternal report of the type of operation performed, there is a possibility for misclassification. However, the results were consistent in a sensitivity analysis evaluating all mothers claiming to have had general anesthetic exposure with an expected bias toward a null effect. Misclassification could also occur if mothers were exposed to anesthesia after the 34-week GA interview and were therefore classified as unexposed. This would bias our results toward a null effect but would have a limited influence on our results as this likely only occurred in a very small number of women. Fourth, since the CBCL is reported by the parent, it may be at risk of subjectivity. Since behavioral and emotional issues may be seen in settings such as home or school, but may not manifest in the structured setting of a neuropsychological evaluation, parent-reported behavioral measures are standard and validated components of neuropsychological evaluations.[28] Fifth, multiple comparisons adjustment was performed to reduce the risk of making a type I error (finding a difference when there should not be one); however, by applying specific adjustment methods, particularly to outcomes that may be correlated, a Type II error (not finding a difference when there should be one) could also be made. In this study, we evaluated 6 different tests with a total of 11 different scores and subscores, many of which were highly correlated. We applied a Bonferroni adjustment and assumed 6 independent tests. However, there is no consensus on the optimal method for multiple comparisons adjustment, and the method used in the current study may be at risk for either a type I or type II error. Finally, like any observational studies, there may be baseline differences in exposed and unexposed children and mothers. No differences in postnatal surgery and resource utilization in exposed and unexposed children were found, but mothers who received general anesthesia during pregnancy reported lower levels of education and higher rates of smoking and alcohol use during pregnancy. While some factors were accounted for, in any observational study, there is a potential risk of unmeasured confounding in all of our analyses.

In 2016, the FDA issued a warning about the neurodevelopmental effects of anesthetic drugs in "children younger than 3 years or in pregnant women during their third trimester." This warning, however, was not based on empirical evidence from studies of

prenatally exposed children, but rather on data from studies of postnatally exposed children and models of prenatal anesthetic exposure in animals.[29,30] The long-term effects of prenatal anesthetic exposure in children are difficult to assess because they require identifying women with exposure during pregnancy followed by a long duration of follow-up in the children. Prenatal exposure, while relatively infrequent, has important implications for understanding a potential mechanism of anesthetic neurotoxicity. The gestational period in humans is characterized by high fetal sensitivity to neurotoxic agents.[9] While some neurotoxicants only have an effect on children during specific windows of gestation, others have an impact throughout all 3 trimesters, as well as into the postnatal period.[9] In preclinical studies, prenatal anesthetic exposure of rodents equivalent to prenatal human exposure in the second trimester, and prenatal exposure of nonhuman primates equivalent to prenatal human exposure in the third trimester have also resulted in neuronal cell death and behavioral changes.[29,30] At this time, the FDA only warns against prenatal anesthetic exposures in the third trimester. Given that the potential mechanism behind anesthetic neurotoxicity and the exact window of vulnerability in children remain unclear, this recommendation is based more on a lack of published studies evaluating neurotoxicity after first and second trimester exposures as opposed to convincing evidence of the safety of anesthetic exposures in children during those time periods.

In this study, prenatal anesthetic exposures were associated with worse long-term behavioral scores in children. While these results provide evidence in support of the FDA warning against prenatal exposure, they also question whether only children exposed in the third trimester may be at risk of long-term neurodevelopmental effects. Despite these findings, it is important to note that since this study has limitations, and delaying or avoiding necessary surgery in pregnant woman can have significant detrimental effects on the mother and the fetus, no changes to clinical practice should be made until these results are confirmed by other similar studies. ■

ACKNOWLEDGMENTS

We acknowledge the Raine Study team for cohort coordination and data collection. Sincere thanks are extended to all Raine Study participants and their families, as this research could not have been conducted without their participation. We are also grateful to Professor Peter D. Sly, MBBS, FRACP, MD, DSc, Deputy Director, Queensland Children's Medical Research Institute (Brisbane, Australia); Jenny Mountain, MClinEpi, Study Manager, the Raine Study Team, University of Western Australia (Perth, Australia); and Huong Le, Data Officer, the Raine Study Team, University of Western Australia (Perth, Australia) for their help in acquiring the data for the manuscript. We also acknowledge Nolan McDonnell, BHB, MBChB, FANZCA, MClinRes, Staff Specialist, King Edward Memorial Hospital (Perth, Australia) and Michael Paech, MBBS, DRCOG, FRCA, FANZA, FFPMANZCA, FRANZCOG(Hon), University of Western Australia (Perth, Australia) for their advice regarding anesthetic exposure in pregnant women in Western Australia.

DISCLOSURES

Name: Caleb Ing, MD, MS.
Contribution: This author helped to conceive and design the study, acquire the data, analyze and interpret the data, draft and revise the manuscript, approve the final manuscript, and agreed to be accountable for all aspects of the work.
Name: Ruth Landau, MD.
Contribution: This author helped to design the study, interpret the data, revise the manuscript, approve the final manuscript, and agreed to be accountable for all aspects of the work.
Name: David DeStephano, MPH.
Contribution: This author helped to analyze and interpret the data, revise the manuscript, approve the final manuscript, and agreed to be accountable for all aspects of the work.
Name: Caleb Miles, PhD.
Contribution: This author helped to analyze and interpret the data, revise the manuscript, approve the final manuscript, and agreed to be accountable for all aspects of the work.
Name: Britta S. von Ungern-Sternberg, MD, PhD.
Contribution: This author helped to design the study, interpret the data, revise the manuscript, approve the final manuscript, and agreed to be accountable for all aspects of the work.
Name: Guohua Li, MD, DrPH.
Contribution: This author helped to conceive and design the study, interpret the data, revise the manuscript, approve the final manuscript, and agreed to be accountable for all aspects of the work.
Name: Andrew J. O. Whitehouse, PhD.
Contribution: This author helped to conceive and design the study, acquire the data, interpret the data, revise the manuscript, approve the final manuscript, and agreed to be accountable for all aspects of the work.
This manuscript was handled by: James A. DiNardo, MD, FAAP.

REFERENCES

1. Jevtovic-Todorovic V, Hartman RE, Izumi Y, et al. Early exposure to common anesthetic agents causes widespread neurodegeneration in the developing rat brain and persistent learning deficits. *J Neurosci.* 2003;23:876–882.
2. Loepke AW, Soriano SG. An assessment of the effects of general anesthetics on developing brain structure and neurocognitive function. *Anesth Analg.* 2008;106:1681–1707.
3. Vutskits L, Xie Z. Lasting impact of general anaesthesia on the brain: mechanisms and relevance. *Nat Rev Neurosci.* 2016;17:705–717.
4. Flick RP, Katusic SK, Colligan RC, et al. Cognitive and behavioral outcomes after early exposure to anesthesia and surgery. *Pediatrics.* 2011;128:e1053–e1061.
5. Ing C, DiMaggio C, Whitehouse A, et al. Long-term differences in language and cognitive function after childhood exposure to anesthesia. *Pediatrics.* 2012;130:e476–e485.
6. Wilder RT, Flick RP, Sprung J, et al. Early exposure to anesthesia and learning disabilities in a population-based birth cohort. *Anesthesiology.* 2009;110:796–804.
7. FDA Drug Safety Communication. FDA review results in new warnings about using general anesthetics and

sedation drugs in young children and pregnant women. 2016. Available at: http://www.fda.gov/Drugs/DrugSafety/ucm532356.htm?source=govdelivery&utm_medium=email&utm_source=govdelivery. Accessed January 3, 2017.
8. Flood P, Rollins MD. Anesthesia for obstetrics. In: Miller RD, ed. *Miller's Anesthesia*. 8th ed. Elsevier/Saunders, 2015:2328–2359.
9. Heyer DB, Meredith RM. Environmental toxicology: sensitive periods of development and neurodevelopmental disorders. *Neurotoxicology.* 2017;58:23–41.
10. Semple BD, Blomgren K, Gimlin K, Ferriero DM, Noble-Haeusslein LJ. Brain development in rodents and humans: identifying benchmarks of maturation and vulnerability to injury across species. *Prog Neurobiol.* 2013;106–107:1–16.
11. The Raine Study. 2011. Accessed November 8, 2018. Available at: http://www.rainestudy.org.au/.
12. Raven J, Court J, Raven J. *Manual for Raven's Progressive Matrices and Vocabulary Scales-Section 2: Coloured Progressive Matrices*. Oxford Psychologists Press; 1990.
13. Smith A. *Symbol Digit Modalities Test*. Western Psychological Services; 1973.
14. McCarron LT. *MAND McCarron Assessment of Neuromuscular Development: Fine and Gross Motor Abilities*. Common Market Press; 1997.
15. Semel E, Wiig E, Secord W. *Clinical Evaluation of Language Fundamentals*. 3rd ed. Psychological Corporation Harcourt Brace Co; 1995.
16. Dunn L, Dunn L, Williams K, Wang J. *Peabody Picture Vocabulary Test III*. American Guidance Services Inc; 1997.
17. Boone KB, Victor TL, Wen J, Razani J, Pontón M. The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Arch Clin Neuropsychol.* 2007;22:355–365.
18. Sun LS, Li G, Miller TL, et al. Association between a single general anesthesia exposure before age 36 months and neurocognitive outcomes in later childhood. *JAMA.* 2016;315:2312–2320.
19. Warner DO, Zaccariello MJ, Katusic SK, et al. Neuropsychological and behavioral outcomes after exposure of young children to procedures requiring general anesthesia: the Mayo Anesthesia Safety in Kids (MASK) study. *Anesthesiology.* 2018;129:89–105.
20. Wilson DJ. The harmonic mean p-value for combining dependent tests. *Proc Natl Acad Sci U S A.* 2019;116:1195–1200.
21. Ing C, Hegarty MK, Perkins JW, et al. Duration of general anaesthetic exposure in early childhood and long-term language and cognitive ability. *Br J Anaesth.* 2017;119:532–540.
22. Sachs A, Guglielminotti J, Miller R, Landau R, Smiley R, Li G. Risk factors and risk stratification for adverse obstetrical outcomes after appendectomy or cholecystectomy during pregnancy. *JAMA Surg.* 2017;152:436–441.
23. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology.* 2009;20:488–495.
24. Mascha EJ, Dalton JE, Kurz A, Saager L. Statistical grand rounds: understanding the mechanism: mediation analysis in randomized and nonrandomized studies. *Anesth Analg.* 2013;117:980–994.
25. Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: methods, interpretation and bias. *Int J Epidemiol.* 2013;42:1511–1519.
26. VanderWeele TJ. A unification of mediation and interaction: a 4-way decomposition. *Epidemiology.* 2014;25:749–761.
27. McCann ME, de Graaff JC, Dorris L, et al; GAS Consortium. Neurodevelopmental outcome at 5 years of age after general anaesthesia or awake-regional anaesthesia in infancy (GAS): an international, multicentre, randomised, controlled equivalence trial. *Lancet.* 2019;393:664–677.
28. Reynolds CR, Kamphaus RW, Vannest KJ. Behavior Assessment System for Children (BASC). In: Kreutzer JS, Caplan B, DeLuca J, eds. *Encyclopedia of Clinical Neuropsychology*. Springer, 2011:366–370.
29. Zhao T, Li Y, Wei W, Savage S, Zhou L, Ma D. Ketamine administered to pregnant rats in the second trimester causes long-lasting behavioral disorders in offspring. *Neurobiol Dis.* 2014;68:145–155.
30. Creeley CE, Dikranian KT, Dissen GA, Back SA, Olney JW, Brambrink AM. Isoflurane-induced apoptosis of neurons and oligodendrocytes in the fetal rhesus macaque brain. *Anesthesiology.* 2014;120:626–638.