

Prospectively assessed neurodevelopmental outcomes in studies of anaesthetic neurotoxicity in children: a systematic review and meta-analysis

Caleb Ing^{1,2,*}, William M. Jackson¹, Michael J. Zaccariello³, Terry E. Goldberg⁴, Mary-Ellen McCann⁵, Anneke Grobler^{6,7}, Andrew Davidson^{7,8}, Lena Sun⁹, Guohua Li^{1,2} and David O. Warner¹⁰

¹Department of Anesthesiology, Columbia University Vagelos College of Physicians and Surgeons, New York, NY, USA, ²Department of Epidemiology, Mailman School of Public Health, New York, NY, USA, ³Department of Psychiatry and Psychology, Mayo Clinic, Rochester, MN, USA, ⁴Department of Psychiatry and Anesthesiology, Columbia University Vagelos College of Physicians and Surgeons, New York, NY, USA, ⁵Department of Anesthesiology, Critical Care and Pain Medicine, Harvard Medical School, Boston Children's Hospital, Boston, MA, USA, ⁶Department of Paediatrics, University of Melbourne, Melbourne, Australia, ⁷Murdoch Children's Research Institute, Melbourne, Australia, ⁸Department of Anaesthesia and Pain Management, Royal Children's Hospital, Melbourne, Australia, ⁹Departments of Anesthesiology and Pediatrics, Columbia University Vagelos College of Physicians and Surgeons, New York, NY, USA and ¹⁰Department of Anesthesiology and Perioperative Medicine, Mayo Clinic, Rochester, MN, USA

*Corresponding author. E-mail: ci2119@cumc.columbia.edu

Abstract

Background: Whether exposure to a single general anaesthetic (GA) in early childhood causes long-term neurodevelopmental problems remains unclear.

Methods: PubMed/MEDLINE, Embase, CINAHL, Web of Science, and the Cochrane Library were searched from inception to October 2019. Studies evaluating neurodevelopmental outcomes and prospectively enrolling children exposed to a single GA procedure compared with unexposed children were identified. Outcomes common to at least three studies were evaluated using random-effects meta-analyses.

Results: Full-scale intelligence quotient (FSIQ); the parentally reported Child Behavior Checklist (CBCL) total, externalising, and internalising problems scores; and Behavior Rating Inventory of Executive Function (BRIEF) scores were assessed. Of 1644 children identified, 841 who had a single exposure to GA were evaluated. The CBCL problem scores were significantly higher (i.e. worse) in exposed children: mean score difference (CBCL total: 2.3 [95% confidence interval {CI}: 1.0–3.7], $P=0.001$; CBCL externalising: 1.9 [95% CI: 0.7–3.1], $P=0.003$; and CBCL internalising problems: 2.2 [95% CI: 0.9–3.5], $P=0.001$). Differences in BRIEF were not significant after multiple comparison adjustment. Full-scale intelligence quotient was not affected by GA exposure. Secondary analyses evaluating the risk of these scores exceeding predetermined clinical thresholds found that GA exposure was associated with increased risk of CBCL internalising behavioural deficit (risk ratio [RR]: 1.47; 95% CI: 1.08–2.02; $P=0.016$) and impaired BRIEF executive function (RR: 1.68; 95% CI: 1.23–2.30; $P=0.001$).

Conclusions: Combining results of studies utilising prospectively collected outcomes showed that a single GA exposure was associated with statistically significant increases in parent reports of behavioural problems with no difference in general intelligence.

Keywords: anaesthetic neurotoxicity; behavioural deficit; meta-analysis; neurodevelopment; paediatric anaesthesia; systematic review

Received: 13 August 2020; Accepted: 18 October 2020

© 2020 British Journal of Anaesthesia. Published by Elsevier Ltd. All rights reserved.
For Permissions, please email: permissions@elsevier.com

Editor's key points

- Whether a single exposure to general anaesthesia in early childhood causes long-term neurodevelopmental problems is unclear despite incriminating animal evidence.
- A meta-analysis was performed of prospective studies evaluating neurodevelopmental outcomes in children exposed to a single general anaesthetic procedure compared with unexposed children.
- Exposure to general anaesthesia was associated with increases in parental reports of behavioural problems with no difference in general intelligence.
- Further research is needed to evaluate the clinical significance of these differences and to identify potentially vulnerable children.

Exposure of young animals to clinically utilised general anaesthetic drugs produces neurodegeneration and later problems with learning, memory, and behaviour.¹ Clinical studies have also evaluated long-term neurodevelopmental outcomes in children exposed to anaesthesia.^{2–8} However, there is significant variation in the studies with regard to study design, patient populations, and neurodevelopmental outcomes assessed, hampering their interpretation. This heterogeneity stems from challenges, including inability to randomise children needing surgery to a non-anaesthetic control group, difficulty in accounting for underlying comorbid conditions in children needing surgery, and the need to assess outcomes many years after the exposure. As a result, most studies of anaesthetic neurotoxicity are observational and use pre-existing data sets.^{2–5} Prior meta-analyses of these retrospective studies have reported increased neurodevelopmental deficits in anaesthetic-exposed children.^{9,10} However, the underlying studies included outcomes that may not be the most sensitive for evaluating neurodevelopment after anaesthetic exposure, and often lacked clinical data regarding pre-exposure and perioperative factors that could affect neurodevelopment. This lack of clinical data is problematic, as children with major congenital anomalies or intraoperative complications may be included in the anaesthetic-exposed group, potentially biasing the results and complicating interpretation of these studies.

This systematic review and meta-analysis compares long-term neurodevelopmental outcomes of children with and without exposure to a single episode of general anaesthesia during a predefined study period. Although well-conducted RCTs typically provide effect estimates that are less susceptible to bias than non-randomised studies, the inclusion of observational studies that appropriately control for confounding may be beneficial.¹¹ To address the limitations of previously published meta-analyses, only RCTs and non-randomised studies of children who were prospectively enrolled and tested were included. This was done because these studies were designed to include the most sensitive outcomes for evaluating neurodevelopment after anaesthesia exposure and also allowed for review of clinical data in an attempt to minimise confounding.

Methods

This study was approved by the Institutional Review Board at Columbia University Vagelos Medical Center (New York, NY, USA) as exempt from requiring written/informed consent. A systematic review with a meta-analysis was performed adhering to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses and Meta-analysis of Observational Studies in Epidemiology checklists.¹² The review protocol was not registered in an online database.

Study design and search strategy

The systematic review was performed identifying all published studies evaluating cognitive function after exposure to general anaesthesia or surgery in children <18 yr old. The criteria and search strategy searching PubMed/MEDLINE, Embase, CINAHL, Web of Science, and the Cochrane Library were published by Clausen and colleagues¹³ identifying 67 English language studies published before June 16, 2017. In the present study, an update using the same search criteria save for minor differences in formatting of search terms was performed to identify any additional studies published from June 17, 2017 until October 16, 2019 (Supplementary Appendix 1). This methodology of utilising results from a previously published systematic review was performed to allow for a more efficient review of new evidence.¹⁴

Study selection and data extraction

The population of interest was children exposed to a single general anaesthetic during a predefined study period using contemporary general anaesthetic medications and monitoring. Studies with children exposed predominantly to halothane were excluded, as that medication is no longer available in most anaesthetic practices. After broadly identifying all studies evaluating cognitive function after exposure to general anaesthesia or surgery in children, additional criteria were applied to focus on relevant RCTs and non-randomised studies with prospectively collected neurodevelopmental outcomes.

The following are the inclusion criteria:

- Studies must evaluate neurodevelopmental outcomes in children exposed to a single general anaesthetic during a defined study period compared with children unexposed to general anaesthesia.
- Neurodevelopmental outcomes must be prospectively assessed in exposed and unexposed children who are at least school-aged, defined in this study as 5 yr of age or older, to allow for more accurate cognitive assessment.

The following are the exclusion criteria:

- Studies that only evaluated short-term perioperative cognitive outcomes, such as delirium or anxiety, were not included.
- Studies focusing specifically on children with major chronic conditions (congenital cardiac or other major congenital condition, extreme prematurity, etc.) were excluded. Children in these studies have significant clinical heterogeneity and may have severe baseline medical issues, complicating interpretation of these studies.

Studies that met these additional criteria were reviewed, and all primary and secondary outcome data were considered.

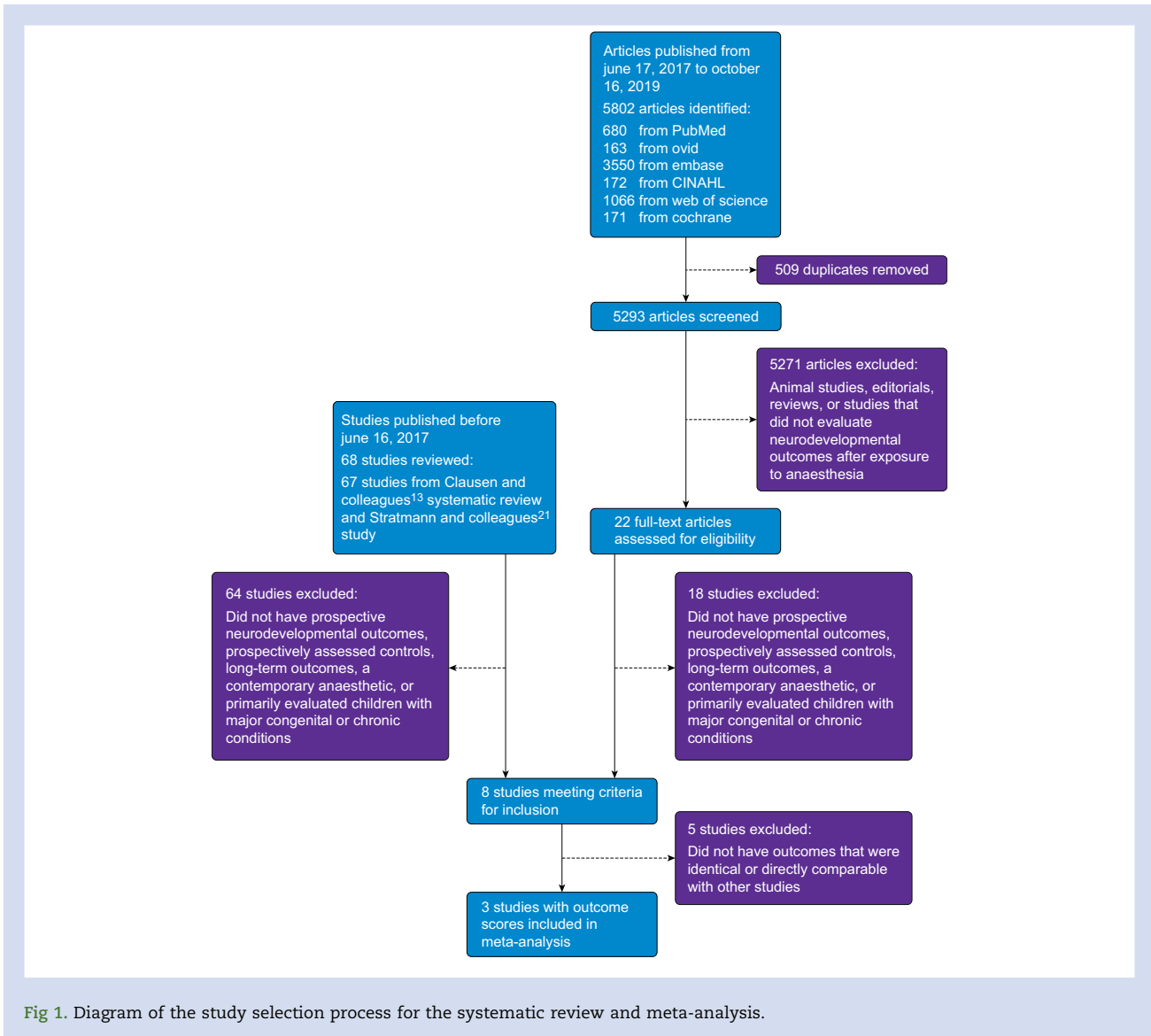


Fig 1. Diagram of the study selection process for the systematic review and meta-analysis.

Any prospectively assessed neurodevelopmental outcome that was measured by the same instruments and common to at least three studies was included in the meta-analysis. Although there is no lower limit of studies needed for a meta-analysis, when only two studies are included, there is a potential for increased risk of Type I error in the setting of heterogeneity between studies.¹⁵ Two reviewers (CI and WMJ) independently assessed titles, abstracts, and full texts for eligibility using Covidence systematic review software (Veritas Health Innovation, Melbourne, VIC, Australia). Conflicts were resolved through consensus and, if necessary, consultation with a third reviewer (DOW). The same reviewers extracted the data from these studies.

Quality assessment and risk of bias

Critical appraisals were conducted using the Cochrane risk-of-bias tool for RCTs and the Risk Of Bias In Non-randomised Studies - of Interventions (ROBINS-I) for non-randomised studies.^{16–18} Response options for the Cochrane risk-of-bias

instrument were 'low' or 'high', whereas response options for the ROBINS-I were 'low', 'moderate', 'serious', or 'critical'. Two reviewers (CI and WMJ) independently assessed each study, with conflicts resolved through consensus or, if necessary, discussion with a third reviewer (GL).

Statistical analysis

The primary analysis evaluated all eligible neurodevelopmental outcome scores as continuous variables. The consistency of outcome scores between studies was evaluated by calculating Cochrane's Q and I^2 statistics for each outcome. Each of the outcome scores was then evaluated using random-effects models. An overall meta-analysis for each outcome was performed by pooling summary data from any outcome used in three or more studies. Publication bias could not be evaluated because of the limited number of available studies.¹⁹ All analyses were performed using Comprehensive Meta-Analysis software, version 3 (Biostat, Englewood, NJ, USA) and figures generated using GraphPad Prism version 8.1.0

Table 1 Neurodevelopmental outcomes evaluated in the eligible studies.

Measure	Study									
	Sun and colleagues (PANDA) ⁷	Warner and colleagues (MASK) ⁸	McCann and colleagues (GAS) ⁶	Taghon and colleagues ²²	Bakri and colleagues ²³	Khochfe and colleagues ²⁴	Warner and colleagues (OTB Study) ²⁵	Zhang and colleagues ²⁶	*Stratmann and colleagues ²¹	
BRIEF Global Executive Composite	✓	✓	✓							
CBCL Externalizing score	✓	✓	✓							
CBCL Internalizing score	✓	✓	✓							
CBCL Total Problems	✓	✓	✓						✓	
CPT-II number commissions	✓	✓								
DKEFS Trail Making Test condition 1	✓	✓								
DKEFS Trail Making Test condition 2	✓	✓								
DKEFS Trail Making Test condition 3	✓	✓								
DKEFS Trail Making Test condition 4	✓	✓								
DKEFS Trail Making Test condition 5	✓	✓								
Grooved Pegboard dominant hand	✓	✓								
Grooved Pegboard nondominant hand	✓	✓								
WASI Full Scale IQ	✓	✓							✓	
WASI Matrix Reasoning	✓	✓								
WASI Vocabulary	✓	✓								
ABAS-II General Adaptive Composite score	✓		✓							
NEPSY-II Speeded naming	✓		✓							
NEPSY-II Word Generation	✓		✓							
ABAS-II Conceptual Composite score	✓									
ABAS-II Practical Composite score	✓									
ABAS-II Social Composite score	✓									
CVLT-C Total Trials 1-5 (Verbal memory)	✓									
NEPSY-II Comprehension of instructions	✓									
NEPSY-II Delayed Memory for Faces	✓									
NEPSY-II Memory for Faces	✓									
WASI Block design	✓									
WASI Performance IQ	✓								✓	
WASI Similarities	✓									
WASI Verbal IQ	✓								✓	
WISC-IV Coding (Processing Speed)	✓									
WISC-IV Digit Span	✓									
Beery: Motor Coordination		✓								
Beery: Visual Motor Integration		✓								
Beery: Visual perception		✓								
Boston Naming Test		✓								
CBCL ADHD Problems		✓								
CLDQ Math Scale		✓								
CLDQ Reading Scale		✓								
CPT-II Detectability		✓								
CPT-II Hit Reaction Time		✓								
CPT-II Hit Reaction Time std error		✓								
CPT-II number omissions		✓								
CPT-II Variability		✓								
CTOPP: Rapid Naming Composite		✓								

Continued

Table 1 Continued

Measure	Study									
	Sun and colleagues (PANDA) ⁷	Warner and colleagues (MASK) ⁸	McCann and colleagues (GAS) ⁶	Taghon and colleagues ²²	Bakri and colleagues ²³	Khochfe and colleagues ²⁴	Warner and colleagues (OTB Study) ²⁵	Zhang and colleagues ²⁶	*Stratmann and colleagues ²¹	
DKEFS Expressive language composite		✓								
DKEFS Tower test Total Achievement Score		✓								
DKEFS Verbal Fluency: Category fluency		✓								
Grooved Pegboard Fine motor composite		✓								
Judgment of Line Orientation		✓								
Visual-spatial abilities composite		✓								
WASI Matrix Vocab		✓								
WCST: Perseverative errors		✓								
WCST: Perseverative responses		✓								
WRAML-2 Attention/concentration Index		✓								
WRAML-2 Delayed verbal recall composite		✓								
WRAML-2 Design Memory subtest		✓								
WRAML-2 Design Recognition subtest		✓								
WRAML-2 Story Memory Delay Recall		✓								
WRAML-2 Story Memory Recognition		✓								
WRAML-2 Verbal Learning Delay Recall		✓								
WRAML-2 Verbal learning Recognition		✓								
WRAML-2 Verbal memory index		✓								
WRAML-2 Verbal recognition composite		✓								
CMS Numbers			✓							
CMS Word Lists I			✓							
CMS Word Lists II			✓							
NEPSY-II Affect Recognition			✓							
NEPSY-II Auditory Attention			✓							
NEPSY-II Design Copy			✓							
NEPSY-II Fingertip tapping repetitions			✓							
NEPSY-II Fingertip tapping sequences			✓							
NEPSY-II Inhibition			✓							
NEPSY-II Memory for Names			✓							
NEPSY-II Sentence Repetition			✓							
NEPSY-II Statue			✓							
NEPSY-II Theory of Mind			✓							
WIAT-II Numerical Composite			✓							
WIAT-II Spelling			✓							
WIAT-II Word reading			✓							
WPPSI-III FSIQ			✓							
WPPSI-III Performance IQ			✓							
WPPSI-III Processing Speed score			✓							
WPPSI-III Verbal IQ			✓							
Go/no go task				✓						
CBCL 1 1/2 - 5 years					✓					
DSM 4th edition (3 point scale)					✓					
Eyberg Child Behaviour Inventory						✓				
OTB							✓			

Continued

Table 1 Continued

Measure	Study
RSPM-IQ	Sun and colleagues (PANDA) ⁷
Color recollection	Warner and colleagues (MASK) ⁸
Spatial recollection	McCann and colleagues (GAS) ⁶
Color familiarity	Taghon and colleagues ²²
Spatial familiarity	Bakri and colleagues ²³
	Khochfe and colleagues ²⁴
	Warner and colleagues (OTB Study) ²⁵
	Zhang and colleagues ²⁶
	*Stratmann and colleagues ²¹

*Study included in sensitivity analysis. ABAS-II, Adaptive Behavior Assessment System, Second Edition; ADHD, attention deficit hyperactivity disorder; BRIEF, Behavior Rating Inventory of Executive Functions; CBCL, Child Behavior Checklist; CLDQ, Colorado Learning Difficulties Questionnaire; CMS, Children's Memory Scale; CPT-II, Continuous Performance Test II; CTOPP, Comprehensive Test of Phonological Processing; CVLT-C, California Verbal Learning Test—Children; DKEFS, Delis-Kaplan Executive Function System; DSM, Diagnostic and Statistical Manual of Mental Disorders; IQ, intelligence quotient; NEPSY-II, Developmental Neuropsychological Assessment, Second Edition; OTB, Operant Test Battery; RSPM, Raven's Standard Progressive Matrices; WASI, Wechsler Abbreviated Scale of Intelligence; WCST, Wisconsin Card Sort Test; WIAT-II, Wechsler Individual Achievement Test, second edition; WISC-IV, Wechsler Intelligence Scale for Children—Fourth Edition; WPPSI-III, Wechsler Preschool and Primary Scale of Intelligence, third edition; WRAML-2, Wide Range Assessment of Memory and Learning, second edition.

(GraphPad Software, San Diego, CA, USA). For the primary analysis, Holm–Bonferroni adjustment was used for multiple comparisons.²⁰

Secondary analysis

A secondary analysis was performed evaluating the parentally reported outcomes after dichotomising the scores using pre-specified clinical cut-offs. For Child Behavior Checklist (CBCL) Total Problems, Externalizing Problems, and Internalizing Problems or Behavior Rating Inventory of Executive Function (BRIEF), scores surpassing a threshold of >60 were >1 standard deviation from the standardised mean score and considered to be a clinically significant deficit, as used in prior studies.^{7,8} Additional data were requested from study authors as necessary to perform this analysis.

Sensitivity analysis

After a review of all published paediatric studies prospectively evaluating cognitive function after exposure to general anaesthesia, one identified study using included outcomes did not independently report scores for children with single vs multiple anaesthetic exposures.²¹ The authors were contacted to obtain scores for the children with single exposures, but the data could not be retrieved. The results from this study, however, were included in a separate sensitivity analysis.

Subgroup analysis of parent-reported outcomes based on blinding status

A criticism of studies using parentally reported outcome measures is that knowledge of their children's exposure to general anaesthesia could bias parents to give their children worse scores if parents believe that general anaesthesia is detrimental. A subgroup analysis was therefore performed using data from the General Anaesthesia or Awake-regional Anaesthesia in Infancy (GAS) trial, in which about half the parents reported being blinded from knowing whether their child received general anaesthesia or regional anaesthesia.⁶ In this analysis, the parentally reported outcome scores were stratified by blinding status to determine if group differences based on type of anaesthetic used were primarily reported by unblinded parents, with blinded parents reporting no difference. These data were evaluated in multiple ways, including as per protocol or intention to treat, and also used multiple imputation as per the original GAS trial analysis.

Results

The systematic review identified 5293 studies published between June 17, 2017 and October 16, 2019 after removal of duplicates. Of these, 22 studies evaluated neurodevelopmental outcomes after exposure to general anaesthesia or surgery in children. Clausen and colleagues¹³ performed an extensive review, which excluded all animal studies. As a result, one notable study by Stratmann and colleagues²¹ was missed, likely because it included data from animals and humans. With inclusion of that additional study, there was a total of 68 studies of neurodevelopmental outcomes after exposure to general anaesthesia or surgery in children before June 16, 2017 and 22 studies from June 17, 2017 until October 16, 2019, for a total of 90 studies of anaesthetic neurotoxicity (Fig. 1). After applying additional inclusion and exclusion criteria to identify

studies that assessed prospectively collected neurodevelopmental outcomes in children exposed to a contemporary general anaesthetic, a total of eight studies remained^{6–8,21–26} (Supplementary Table 1).

Outcomes evaluated

In reviewing the eight eligible studies, five outcome scores were identified that were reported in at least three studies: CBCL total problems, externalising problems, and internalising problems scores; the BRIEF Global Executive Composite score, and full-scale intelligence quotient (FSIQ) (Table 1). For the FSIQ, two of the studies used the Wechsler Abbreviated Scale of Intelligence (WASI) to assess FSIQ,^{7,8} whereas one study used the Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III).⁶ For the BRIEF Global Executive Composite, one of the studies used the preschool version.⁶ For the FSIQ and BRIEF, the scores used were appropriate for the age of the study population. Given the similarities between the FSIQ as measured by the WASI vs the WPPSI-III, and Global Executive Composite score as measured by the BRIEF vs the BRIEF preschool version, they were considered to be the same instruments for the purposes of the meta-analysis. In interpreting the parentally reported outcome scores, higher scores represent worse behaviour (CBCL externalising problems), increased emotional distress (CBCL internalising problems), and more impaired executive function (BRIEF Global Executive Composite).

Study characteristics

Three studies contributed data for the FSIQ, BRIEF, and CBCL outcome scores. The GAS trial enrolled children scheduled for inguinal hernia repair (mean age ~70 days), and randomised them to receive either general anaesthesia with sevoflurane or regional anaesthesia with spinal or caudal blocks with neurodevelopmental evaluation at 5 yr of age.⁶ The two other studies relied on an ‘ambi-directional’ observational approach, with children old enough to undergo prospective neuropsychological testing retrospectively identified as having been exposed to surgery and anaesthesia at ≤ 3 yr of age. The Pediatric Anesthesia Neurodevelopment Assessment (PANDA) study⁷ included siblings discordant for exposure to hernia surgery with neurodevelopmental evaluation at 8–15 yr of age, and the Mayo Anesthesia Safety in Kids (MASK) study⁸ included children undergoing a variety of surgical procedures with children singly or multiply exposed to general anaesthesia before age 3 yr propensity matched to unexposed children with neurodevelopmental evaluation at 8–12 or 15–20 yr.

In all three studies, these outcomes were evaluated as continuous variables and presented in each publication as mean differences between exposed and unexposed children with 95% confidence intervals (CIs) around those differences.^{6–8} For the GAS trial, the results were evaluated in several different ways. For the purposes of this meta-analysis, the results from the multiple imputation per protocol analysis were used.⁶ For the MASK study, only data from children singly exposed to anaesthesia were included.⁸

Risk of bias

All studies were at risk of bias with the GAS trial at risk because of incomplete blinding and loss to follow up, whereas

the PANDA and MASK studies were at risk because of confounding and selection bias (Supplementary Tables 2 and 3).

Outcomes after general anaesthetic exposure

Of 1644 children included, outcome data were available for 837–841 children exposed to general anaesthesia and 799–803 children unexposed to general anaesthesia depending on the outcome score (Fig. 2). Regarding heterogeneity in the outcomes between studies, the FSIQ and CBCL scores were consistent between all three studies, with $I^2=0\%$, but substantial between-study heterogeneity amongst the BRIEF scores with $I^2=69\%$ was seen. After pooling summary data from all three studies, the difference in mean scores between those exposed and unexposed to general anaesthesia was -0.2 (95% CI: -1.7 to 1.3), $P=0.79$ for FSIQ; 2.3 (95% CI: 1.0 – 3.7), $P=0.001$ for CBCL total problems, 1.9 (95% CI: 0.7 – 3.1), $P=0.003$ for CBCL externalising problems, and 2.2 (95% CI: 0.9 – 3.5), $P=0.001$ for CBCL internalising problems; and 2.6 (95% CI: 0.1 – 5.0), $P=0.036$ for BRIEF scores. After adjustment for multiple comparisons, the differences in all CBCL scores remained statistically significant, but the differences in BRIEF scores were no longer significant.

In the secondary analysis, the increased risk of the score exceeding a predetermined threshold for clinical deficit was evaluated, with a single exposure associated with an increased risk of subsequent CBCL internalising behavioural deficit (risk ratio [RR]: 1.47 ; 95% CI: 1.08 – 2.02 ; $P=0.016$) and impaired executive function (RR: 1.68 ; 95% CI: 1.23 – 2.30 ; $P=0.001$) (Fig. 3).

In a sensitivity analysis, data from the Stratmann and colleagues²¹ paper were added, which included 28 exposed and 28 unexposed children. Of the exposed children, 64% ($n=18$) received a single anaesthetic. Inclusion of the FSIQ and CBCL total scores from this study did not substantially alter the primary results (Supplementary Fig. 1).

Parent-reported outcomes according to blinding status

In the GAS trial, blinding to treatment assignment was achieved in 51% ($n=256$) of parents, whereas the remainder reported knowing the type of anaesthetic their child received. The proportion of blinded parents in the regional and general anaesthesia groups was also similar, with parental blinding in 51% ($n=105$) of children receiving regional anaesthesia and 49% ($n=118$) of children receiving general anaesthesia. In blinded parents, the general-anaesthetic-exposed children had mean CBCL scores that were between 0.4 and 2.5 points higher, and BRIEF scores that were 1.3–1.6 points higher than the mean scores of children with a regional anaesthetic (Table 2). In unblinded parents, the general-anaesthetic-exposed children had mean CBCL scores that were between 0.5 and 1.8 points higher, and BRIEF scores that were 2.2–2.9 points higher than the mean scores of children with a regional anaesthetic.

Discussion

We evaluated all prospectively assessed neurodevelopmental outcome data from prospectively designed studies comparing general-anaesthetic-exposed children with children not exposed to general anaesthesia up to October 2019. Studies included children exposed to anaesthesia at under 1 yr of age⁶ and under 3 yr of age.^{7,8} Compared with children who did not receive general anaesthesia, children exposed to a single

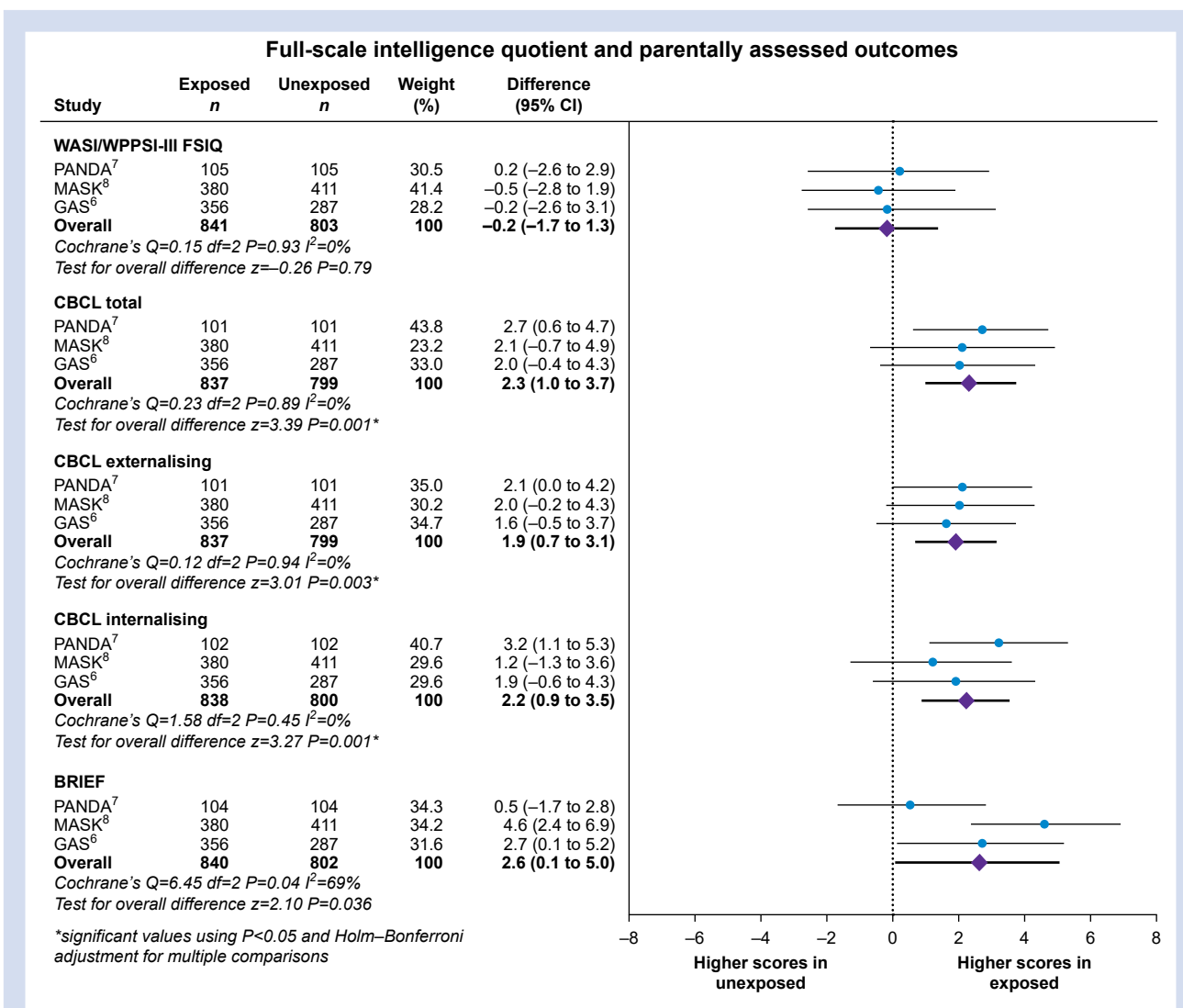


Fig 2. Full-scale intelligence quotient and parent-reported outcome scores in children with a single exposure to general anaesthesia vs no exposure to general anaesthesia. BRIEF, Behavior Rating Inventory of Executive Function; CBCL, Child Behavior Checklist; CI, confidence interval; FSIQ, full-scale intelligence quotient; GAS, General Anaesthesia or Awake-regional Anaesthesia in Infancy; MASK, Mayo Anesthesia Safety in Kids; PANDA, Pediatric Anesthesia Neurodevelopment Assessment; WASI, Wechsler Abbreviated Scale of Intelligence; WPPSI-III, Wechsler Preschool and Primary Scale of Intelligence, Third Edition.

general anaesthetic had mean score differences in CBCL and BRIEF from 1.9 to 2.6 points worse than those in unexposed children, which, in secondary analyses, corresponded to a 47% increased risk of an internalising behavioural deficit and a 68% increased risk of impaired executive function. No significant differences in FSIQ were found.

Parent reports of behaviour and executive function are standard and useful components of clinical neuropsychological evaluations, as behaviour or emotional issues may not manifest in the structured setting of a neuropsychological assessment, but are evident in other settings, such as home or school.²⁷ The finding of more problems in children exposed to general anaesthesia is consistent with studies of non-human primates exposed to anaesthetics that reported behavioural problems after early anaesthetic exposure.²⁸ It is also

consistent with some retrospective analyses suggesting that children exposed to anaesthesia may be more likely to develop later behavioural problems, such as attention deficit hyperactivity disorder.^{3,29–31} Thus, measures of behaviour may be of particular interest in the search for a potential phenotype associated with anaesthesia exposure. One criticism of parent reports is that if parents know that their child was exposed to general anaesthesia, they may be biased towards reporting more problems. The results of the GAS trial, an RCT with blinding to treatment assignment maintained for about half of participants, provide insight into this possibility. If bias was the factor causing more problems to be reported in anaesthetic-exposed children, then worse scores might primarily be seen in unblinded parents, with blinded parents reporting no difference. Although the small sample size

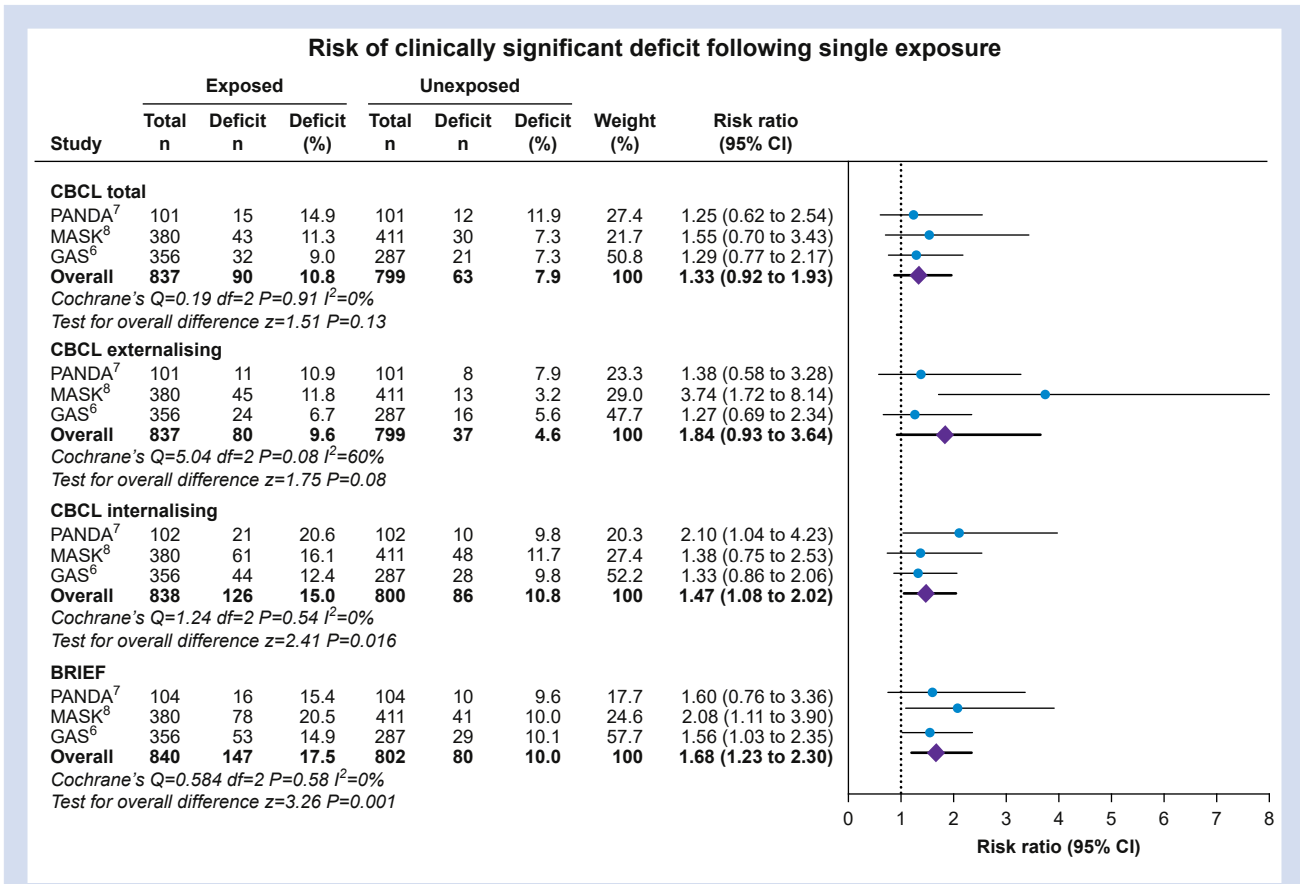


Fig 3. Clinically significant deficit in parent-reported outcome scores in children with a single exposure to general anaesthesia vs no exposure to general anaesthesia. BRIEF, Behavior Rating Inventory of Executive Function; CBCL, Child Behavior Checklist; CI, confidence interval; GAS, General Anaesthesia or Awake-regional Anaesthesia in Infancy; MASK, Mayo Anesthesia Safety in Kids; PANDA, Pediatric Anesthesia Neurodevelopment Assessment.

precluded a formal analysis, there was little evidence that blinding status was consistently associated with differences between parent-reported scores in general and regional anaesthesia groups, as similar results were seen in both groups.

The clinical relevance of these score differences (~0.2 standard deviations or 2 points on CBCL and BRIEF) identified in children undergoing short single procedures is unclear, particularly on an individual level. However, on a broader scale, if these worse behavioural scores are actually caused by anaesthesia, given that 500 000 to 1 million children are exposed to anaesthesia in early childhood each year in the USA alone, these differences may have increased importance on a population level.^{32,33} To further evaluate the clinical significance of these score differences, we evaluated the percentage of children crossing a predetermined clinical threshold for deficit. A single general anaesthetic exposure significantly increased the risk of developing internalising behavioural problems by 47% and impaired executive function by 68%. Whether these score differences represent a shift for the entire population of general-anaesthetic-exposed children or only a small group of vulnerable children is unclear. If an entire normal distribution is shifted to the right by 0.2 standard deviations, the percentage of children with scores above 1 standard deviation from the mean (equivalent to >60 on the

CBCL or BRIEF) would increase by approximately 34%. This is similar to the increased risk found in this study, and therefore could suggest that all children were affected. However, given the uncertainty around these estimates, it is also possible that this shift in scores represents a more severe impact on a vulnerable group of children in combination with some children who were unaffected.

A limitation of this analysis is the heterogeneity amongst the included studies in design, control condition examined, and analytical methods used. The age at evaluation also differed, with children from the GAS study evaluated at age 5 yr, whereas children from the MASK and PANDA studies were evaluated at 8–20 yr of age. Although meta-analyses commonly include studies with methodological differences, heterogeneity poses a threat to the validity of combining the results of these studies. Heterogeneity was also seen between studies in BRIEF scores despite consistency in FSIQ and CBCL scores. Another limitation is that children were characterised as being exposed to general anaesthesia based on exposure during the assessment period defined in each study. Although balanced between exposed and unexposed groups, some children in all of these studies were exposed to general anaesthesia after the study assessment period, which could bias against finding differences if these later exposures affected neurodevelopment.

Table 2 Comparison of general anaesthetic vs regional anaesthetic groups stratified by blinding status.

	GA group		RA group		Difference in GA–RA (95% CI)*
	N	Mean (sd)	n	Mean (sd)	
Blinded parents					
CBCL total					
PP multiple imputation	132	47 (12.6)	100	45.4 (11.9)	2.1 (–1.1 to 5.4)
PP complete case	129	46.9 (12.2)	97	45.4 (11.4)	2 (–1.2 to 5.1)
ITT multiple imputation	132	47 (12.6)	124	45.7 (12.6)	1.8 (–1.4 to 4.9)
ITT complete case	129	46.9 (12.2)	120	45.7 (12)	1.6 (–1.5 to 4.6)
CBCL externalising					
PP multiple imputation	132	45.6 (11.8)	100	45.1 (11.3)	0.8 (–2.2 to 3.9)
PP complete case	129	45.5 (11.5)	97	45 (10.9)	0.7 (–2.4 to 3.7)
ITT multiple imputation	132	45.6 (11.8)	124	45.2 (11.9)	0.5 (–2.4 to 3.5)
ITT complete case	129	45.5 (11.5)	120	45.2 (11.4)	0.4 (–2.5 to 3.3)
CBCL internalising					
PP multiple imputation	132	48.5 (12.8)	100	46.5 (12.4)	2.5 (–0.9 to 5.9)
PP complete case	129	48.4 (12.4)	97	46.5 (11.9)	2.4 (–0.9 to 5.7)
ITT multiple imputation	132	48.5 (12.8)	124	46.7 (12.8)	2.3 (–0.9 to 5.5)
ITT complete case	129	48.4 (12.4)	120	46.7 (12.2)	2.2 (–1 to 5.3)
BRIEF					
PP multiple imputation	132	51.1 (14.4)	100	49.9 (14.4)	1.5 (–2.3 to 5.3)
PP complete case	116	51 (12.5)	90	49.5 (12.5)	1.6 (–1.9 to 5.1)
ITT multiple imputation	132	51.1 (14.4)	124	50 (14.5)	1.3 (–2.3 to 4.9)
ITT complete case	116	51 (12.5)	112	49.8 (13)	1.3 (–2.1 to 4.7)
Unblinded parents					
CBCL total					
PP multiple imputation	118	46 (12.7)	105	44.2 (12.3)	1.1 (–2.2 to 4.4)
PP complete case	116	45.9 (12.5)	104	44.1 (12.2)	1 (–2.3 to 4.3)
ITT multiple imputation	118	46 (12.7)	131	44.6 (12.5)	0.7 (–2.4 to 3.9)
ITT complete case	116	45.9 (12.5)	129	44.5 (12.3)	0.7 (–2.4 to 3.8)
CBCL externalising					
PP multiple imputation	118	45.7 (12.5)	105	43.3 (10.8)	1.8 (–1.3 to 4.9)
PP complete case	116	45.6 (12.3)	104	43.3 (10.7)	1.7 (–1.4 to 4.8)
ITT multiple imputation	118	45.7 (12.5)	131	43.9 (11.8)	1.2 (–1.8 to 4.2)
ITT complete case	116	45.6 (12.3)	129	43.9 (11.6)	1.1 (–1.8 to 4.1)
CBCL internalising					
PP multiple imputation	118	47.4 (12.6)	105	46.1 (12.9)	0.5 (–2.9 to 3.9)
PP complete case	116	47.3 (12.4)	104	46.1 (12.9)	0.5 (–2.9 to 3.9)
ITT multiple imputation	118	47.4 (12.6)	131	46.1 (12.9)	0.6 (–2.5 to 3.8)
ITT complete case	116	47.3 (12.4)	129	46 (12.7)	0.7 (–2.5 to 3.8)
BRIEF					
PP multiple imputation	118	50.5 (14.6)	105	47.3 (13.5)	2.5 (–1.2 to 6.3)
PP complete case	107	51 (14.1)	95	47.3 (13)	2.9 (–0.9 to 6.7)
ITT multiple imputation	118	50.5 (14.6)	131	47.6 (13.3)	2.2 (–1.3 to 5.7)
ITT complete case	107	51 (14.1)	119	47.8 (12.7)	2.6 (–1 to 6.1)

*Adjusted for gestational age at birth and country. BRIEF, Behavior Rating Inventory of Executive Function; CBCL, Child Behavior Checklist; ITT, intention-to-treat analysis of data; PP, as per-protocol analysis of data.

Conclusions

Combining the results of studies utilising common prospectively collected outcomes shows that a single exposure to general anaesthesia in early childhood was associated with statistically significant increases in parent reports of behavioural problems, but no difference in general intelligence. Although there was heterogeneity in study methodology, interestingly, mean score differences for behavioural problems in the included studies were strikingly similar. Further research is needed to evaluate the clinical significance of these differences and to identify potentially vulnerable children. The limitations of this analysis mean that these are provisional conclusions that require further research for confirmation.

Authors' contributions

Study conception/design: CI, WMJ, DOW

Data acquisition: CI, WMJ, M-EM, AD, LS, DOW

Data analysis, interpretation, or both: CI, WMJ, MJZ, TEG, M-EM, AG, AD, GL, DOW

Writing/revising of paper: all authors

Final approval: all authors

Agreement to be accountable for all aspects of the work: all authors

Acknowledgements

The authors would like to acknowledge the General Anaesthesia or Awake-regional Anaesthesia in Infancy consortium,

and the Mayo Anesthesia Safety in Kids and Pediatric Anesthesia Neurodevelopment Assessment study collaborators for providing additional unpublished data needed to complete the analyses. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Health Research and Quality.

Declarations of interest

The authors declare that they have no conflicts of interest.

Funding

Agency for Healthcare Research and Quality (AHRQ) (R01 HS026493) to CI.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bja.2020.10.022>.

References

- Vutskits L, Xie Z. Lasting impact of general anaesthesia on the brain: mechanisms and relevance. *Nat Rev Neurosci* 2016; **17**: 705–17
- Glatz P, Sandin RH, Pedersen NL, Bonamy AK, Eriksson LI, Granath F. Association of anesthesia and surgery during childhood with long-term academic performance. *JAMA Pediatr* 2017; **171**, e163470
- Ing C, Ma X, Sun M, et al. Exposure to surgery and anesthesia in early childhood and subsequent use of attention deficit hyperactivity disorder medications. *Anesth Analg* 2020; **131**: 723–33
- O'Leary JD, Janus M, Duku E, et al. A population-based study evaluating the association between surgery in early life and child development at primary school entry. *Anesthesiology* 2016; **125**: 272–9
- 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
- McCann ME, de Graaff JC, Dorris L, et al. 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**: 664–77
- 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**: 2312–20
- 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
- Dimaggio C, Sun LS, Ing C, Li G. Pediatric anesthesia and neurodevelopmental impairments: a Bayesian meta-analysis. *J Neurosurg Anesthesiol* 2012; **24**: 376–81
- Zhang H, Du L, Du Z, Jiang H, Han D, Li Q. Association between childhood exposure to single general anesthesia and neurodevelopment: a systematic review and meta-analysis of cohort study. *J Anesth* 2015; **29**: 749–57
- Shrier I, Boivin JF, Steele RJ, et al. Should meta-analyses of interventions include observational studies in addition to randomized controlled trials? A critical examination of underlying principles. *Am J Epidemiol* 2007; **166**: 1203–9
- Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000; **283**: 2008–12
- Clausen NG, Kahler S, Hansen TG. Systematic review of the neurocognitive outcomes used in studies of paediatric anaesthesia neurotoxicity. *Br J Anaesth* 2018; **120**: 1255–73
- Garner P, Hopewell S, Chandler J, et al. When and how to update systematic reviews: consensus and checklist. *BMJ* 2016; **354**: i3507
- Gonnermann A, Framke T, Grosshennig A, Koch A. No solution yet for combining two independent studies in the presence of heterogeneity. *Stat Med* 2015; **34**: 2476–80
- Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011; **343**: d5928
- Sterne JA, Hernan MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016; **355**: i4919
- Sterne JAC, Savovic J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019; **366**: l4898
- Dalton JE, Bolen SD, Mascha EJ. Publication bias: the elephant in the review. *Anesth Analg* 2016; **123**: 812–3
- Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat* 1979; **6**: 65–70
- Stratmann G, Lee J, Sall JW, et al. Effect of general anesthesia in infancy on long-term recognition memory in humans and rats. *Neuropsychopharmacology* 2014; **39**: 2275–87
- Taghon TA, Masunga AN, Small RH, Kashou NH. A comparison of functional magnetic resonance imaging findings in children with and without a history of early exposure to general anesthesia. *Paediatr Anaesth* 2015; **25**: 239–46
- Bakri MH, Ismail EA, Ali MS, Elsedfy GO, Sayed TA, Ibrahim A. Behavioral and emotional effects of repeated general anesthesia in young children. *Saudi J Anaesth* 2015; **9**: 161–6
- Khochfe AR, Rajab M, Ziade F, Naja ZZ, Naja AS, Naja ZM. The effect of regional anaesthesia versus general anaesthesia on behavioural functions in children. *Anaesth Crit Care Pain Med* 2019; **38**: 357–61
- Warner DO, Chelonis JJ, Paule MG, et al. Performance on the operant test battery in young children exposed to procedures requiring general anaesthesia: the MASK study. *Br J Anaesth* 2019; **122**: 470–9
- Zhang Q, Peng Y, Wang Y. Long-duration general anesthesia influences the intelligence of school age children. *BMC Anesthesiol* 2017; **17**: 170
- Reynolds CR, Kamphaus RW, Vannest KJ. Behavior assessment system for children (BASC). In: Kreutzer JS, Caplan B, DeLuca J, editors. *Encyclopedia of clinical neuropsychology*. New York/London 366-370.: Springer; 2011
- Coleman K, Robertson ND, Dissen GA, et al. Isoflurane anesthesia has long-term consequences on motor and behavioral development in infant rhesus macaques. *Anesthesiology* 2017; **126**: 74–84
- Hu D, Flick RP, Zaccariello MJ, et al. Association between exposure of young children to procedures requiring general anesthesia and learning and behavioral outcomes in a population-based birth cohort. *Anesthesiology* 2017; **127**: 227–40

30. Ing C, Sun M, Olfson M, et al. Age at exposure to surgery and anesthesia in children and association with mental disorder diagnosis. *Anesth Analg* 2017; **125**: 1988–98
31. Flick RP, Katusic SK, Colligan RC, et al. Cognitive and behavioral outcomes after early exposure to anesthesia and surgery. *Pediatrics* 2011; **128**: e1053–61
32. Tzong KY, Han S, Roh A, Ing C. Epidemiology of pediatric surgical admissions in US children: data from the HCUP Kids Inpatient Database. *J Neurosurg Anesthesiol* 2012; **24**: 391–5
33. Shi Y, Hu D, Rodgers EL, et al. Epidemiology of general anesthesia prior to age 3 in a population-based birth cohort. *Paediatr Anaesth* 2018; **28**: 513–9

Handling editor: Hugh C Hemmings Jr