Original Article

Neurodevelopmental outcomes after prenatal exposure to anaesthesia for maternal surgery: a propensity-score weighted bidirectional cohort study

T. Bleeser,[1] **S. Devroe,**[2] **N. Lucas,**[3] **T. Debels,**[4] **M. Van de Velde,**[5] **J. Lemiere,**[6] **J. Deprest**[7] and **S. Rex**[5]

1 Resident, 2 Staff Anaesthetist, 5 Clinical Professor, Department of Anaesthesiology, University Hospitals Leuven, Belgium
3 Consultant, Department of Anaesthesia, Northwick Park Hospital, Harrow, UK
4 Medical Student, Faculty of Medicine, KU Leuven, 6 Clinical Professor, Department of Pediatric Hemato-Oncology, 7 Clinical Professor, Department of Obstetrics and Gynaecology, University Hospitals Leuven, Belgium

Summary

Up to 1% of pregnant women undergo anaesthesia for non-obstetric surgery. This study investigated neurodevelopmental outcomes after prenatal anaesthesia for maternal surgery. A bidirectional cohort study of children born between 2001 and 2018 was performed: neurodevelopmental outcomes of children who had received prenatal anaesthesia for maternal surgery were prospectively compared with unexposed children, with exposure status being assessed retrospectively. Children exposed to anaesthesia for obstetric and fetal surgery were excluded. The primary outcome was the global executive composite of the behaviour rating inventory of executive function score. Our secondary outcomes were: total problems; internalising problems and externalising problems derived from the child behaviour checklist; psychiatric diagnoses; and learning disorders. In 90% of exposed children, there was a single mean (SD) antenatal anaesthesia exposure lasting 91 (94) min. There was a broad spectrum of indications, with abdominal surgery being most frequent. Parents of 129 exposed (response rate 68%) and 453 unexposed (response rate 63%) children participated. There were no arguments for non-response bias. After propensity weighting, there were no statistically significant differences in primary outcome, with a weighted mean difference (95%CI) of exposed minus unexposed children of 1.9 (−0.4–4.2), p = 0.10; or any of the secondary outcomes. Sensitivity analyses confirmed the robustness. Exploratory analyses, however, showed significant differences in certain subgroups for the primary outcome, (e.g. for intra-abdominal surgery, exposure duration > 1 h) and some cognitive subdomains (e.g. working memory and attention). This bidirectional cohort study, the largest investigation on the subject to date, has found no evidence in the general population for an association between prenatal exposure to anaesthesia and impaired neurodevelopmental outcomes.

Correspondence to: S. Rex
Email: steffen.rex@uzleuven.be
Accepted: 20 September 2022
Keywords: anaesthesia; child; neurodevelopment; pregnancy; surgery
Presented in part at Euroanaesthesia, Milan, June 2022.

Introduction

Up to 1% of pregnant women require anaesthesia during pregnancy for non-obstetric surgery [1, 2]. Surgery performed during pregnancy is usually undertaken as an emergency, e.g. appendicectomy and surgery for adnexal pathology [1].

A recent meta-analysis of animal studies demonstrated that general anaesthesia during pregnancy can induce neuronal injury in the fetus and impair learning and memory [3]. However, it is still uncertain whether these findings can be translated to the clinical setting [3]. Despite the limitations of this evidence, the US Food and Drug Administration published a warning in 2016 that repeated or prolonged use of general anaesthesia in pregnant women during their third trimester may result in impaired neurodevelopmental outcomes for the exposed children [4].

We hypothesised that in-utero exposure to anaesthesia would be associated with impairment in neurodevelopmental outcomes in children. To test our hypothesis, we performed a bidirectional cohort study, with retrospective identification of children who were prenatally exposed to anaesthesia for maternal non-obstetric surgery. The neurodevelopmental outcomes of these children were compared, using parental questionnaires, with that of unexposed children through a prospective assessment of executive function, psychosocial problems, diagnoses listed in the 5th edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) [5] and learning disorders, in children aged 2–18 y.

Methods

Ethical committee approval was obtained for the study. Informed consent was obtained from parents in writing or online.

This bidirectional cohort study (retrospective assessment of the exposure status and prospective assessment of neurodevelopmental outcomes) was performed at University Hospitals Leuven, a tertiary referral centre offering services for both low- and high-risk patients.

In the present study, the cohort of our previous publication reporting on pregnancy outcomes after non-obstetric surgery was updated [1]. Children aged between 2 and 18 y on 1 June 2020, were eligible. Participants were defined as `exposed´ when their mothers had undergone general or regional anaesthesia during pregnancy for maternal non-obstetric surgery [1]. Mothers of unexposed children did not have such an exposure [1]. Exclusion criteria for both groups were fetal interventions/surgery, obstetric surgery and deceased children [1].

A computer algorithm screened the hospital database for mothers who had undergone anaesthesia within 280 days of delivery [1]. Retrieved data were verified manually. For each exposed child, four unexposed children [6] were identified who were born at the closest date of birth to the exposed child and born to women of the same birth year and parity as the mother of the exposed child [1]. In both groups, in cases of twins/triplets, one sibling was selected by using a computer random number generator. Therefore, in all numbers mentioned in this article, twins/triplets were counted as one observation.

On 29 April 2020, an invitation letter and questionnaire were sent by post to the parents. The parents from whom no response was obtained were reminded by telephone calls or emails up to eight times with intervals of ≥ 1 month until they completed the questionnaire or decided not to participate. The last contact with participants was on 4 August 2021 (see online Supporting Information, Figure S1). Parents could choose to complete the questionnaires online or return them by mail.

Four questionnaires were used (see online Supporting Information, Appendix S1). First, general information and demographic variables were collected. Second, executive function was assessed using the behaviour rating inventory of executive function parent questionnaire (BRIEF), further referred to as the `executive function questionnaire´) [7]. Third, psychosocial problems were evaluated using the child behaviour checklist parent questionnaire (CBC) further referred to as the `psychosocial problem questionnaire´ [8]. Both questionnaires provide normative t-scores. In a representative population for Dutch-speaking children, the average (SD) t-score is 50 (10). A t-score ≥ 60 suggests a child at risk or a clinically-relevant problem [7, 8]. Fourth, for children aged ≥ 6 y, parents were asked whether their child had ever been diagnosed (by a family doctor, psychologist or psychiatrist) with one of the following disorders described in the DSM-5 [5]: attention-deficit/hyperactivity disorder; autistic spectrum disorder; anxiety disorders; schizophrenia; bipolar disorder; depressive disorders; or personality disorders and learning disorders including dyslexia, dysgraphia and dyscalculia.

The primary outcome was the t-score of the global executive composite (further referred to as the `composite score´) of the executive function questionnaire. This composite score is the sum of all items of the executive function questionnaire and represents a child's overall executive function. Secondary outcomes included the t-scores of the psychosocial problem questionnaire total problems, internalising problems and externalising

problems, and the risks for DSM-5 [5] diagnoses, and learning disorders.

All data were obtained from the questionnaires and the hospital database (see online Supporting Information, Appendix S1) and SAS System for Windows (version 9.4, SAS Institute Inc, Cary, NC, USA) was used. The significance level was a priori set at 0.05, and all tests were two-sided.

We estimated a priori whether a clinically-relevant difference could be detected for the primary outcome in our cohort. Based on our previous study [1] and assuming a response rate of 50%, we estimated we would receive completed questionnaires from at least 90 exposed and 360 unexposed children. Assuming that the standard deviation for the t-scores would be comparable to that of a normative population of children (i.e. 10), a two-sided unpaired t-test with an α of 0.05 showed that there is at least 80% power to detect a difference for the primary outcome equal to 3.3. This effect size is comparable with the observations of prospective studies assessing postnatal exposure to anaesthesia [9, 10].

Non-response bias was assessed by comparing all known characteristics of respondents with those of non-respondents and by wave analysis [11, 12] (see online Supporting Information, Appendix S1). If data were missing from questionnaires, we contacted the parents concerned; thereafter, all data were complete.

Propensity scores were used in an inverse probability of treatment weighting approach to reduce bias by confounding (see online Supporting Information, Appendix S1). Briefly, for every child, a weight was calculated to account for known confounders, i.e. birth month of the child; maternal age at birth; sex of the child; parity of the mother; number of fetuses; maternal exposure during pregnancy to radiation; smoking and alcohol consumption; exposure of the child to anaesthesia after birth; university level and income of the parents; owning vs. renting the family residence; geographic origin of the parents; rural/urban residence; and marital status [9, 10, 13–17]. These weights were taken into account when comparing the outcomes of exposed vs. unexposed children by linear (continuous variables) and Poisson (dichotomous variables) regression models.

Data are expressed as (weighted/unweighted mean difference of t-scores or weighted/unweighted absolute risk reduction, 95% confidence interval; p-value). These differences represent exposed minus unexposed children; therefore, positive values indicate a worse outcome for exposed children. Continuous variables were compared with the Student's t-test; categorical variables were compared using Fisher's exact test. No correction for multiple testing was used.

Sensitivity and exploratory analyses (see online Supporting Information, Appendix S1) were performed using inverse probability of treatment weighting unless stated otherwise. No correction for multiple testing was used, so the exploratory analyses need to be considered as merely hypothesis-generating. To illustrate the practical interpretation of the effect size of the primary outcome, the effect size of in utero exposure to anaesthesia was compared with the effect sizes of all 15 confounders taken into account by inverse probability of treatment weighting. When investigating the effect of these 15 confounders on the neurodevelopmental outcomes of the children, weights were calculated using all other 14 confounders and in utero exposure to anaesthesia (see online Supporting Information, Appendix S1 for details).

Results

We invited the parents of 189 exposed and 721 unexposed children, and received responses from 129 exposed (response rate 68.3%) and 453 unexposed (response rate 62.8%) children that could be included in the analysis (Fig. 1).

Respondents and non-respondents differed significantly in only two out of the 24 baseline characteristics (see online Supporting Information, Table S1). According to the wave analysis, the t-scores' estimated bias (caused by non-response) is 0.14 for the composite score of the executive function questionnaire (behaviour rating inventory of executive function, BRIEF) and 0.48 for the psychosocial problem questionnaire (child behaviour checklist, CBC) total problems (both expressed on the scale of t-scores), that is, < 5% of the standard deviation of the normative t-score of our population.

At baseline, in the exposed group, there were significantly more preterm births and children had a lower birth weight. There was a higher incidence of malignant pathology in the mother, and children were more frequently exposed in utero to maternal chemotherapy, radiography, computer tomography scans and smoking. The university level of the father and the combined income of both parents were significantly lower, and the parental divorce rate was significantly higher (online Supporting Information, Table S2). There were no significant differences for other variables. Applying inverse probability of treatment weighting made the baseline differences between the exposed and unexposed groups negligible (see online Supporting Information, Figure S2).

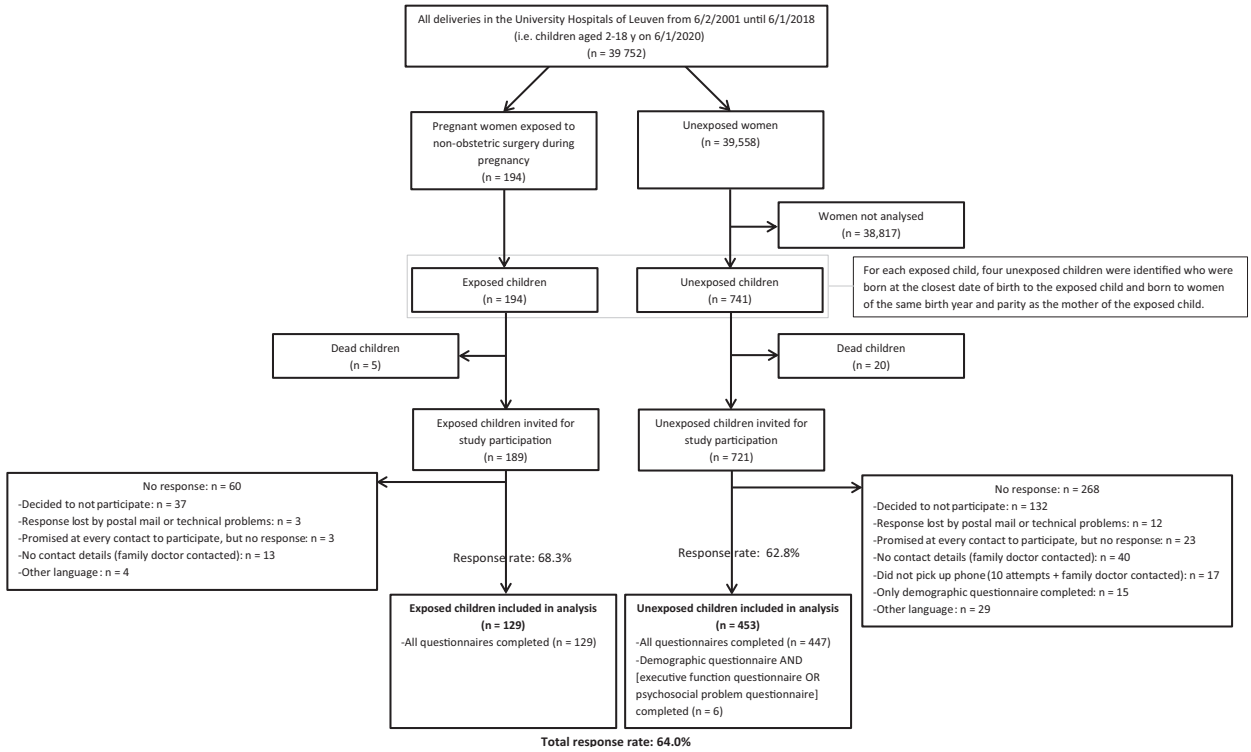

Figure 1 Study flow diagram. Twins/siblings are counted as one observation for all numbers.

In 90% of exposed children, there was a single mean (SD) antenatal anaesthesia exposure lasting 91(94) min (online Supporting Information, Table S2). In 86% of these children, the mothers were exposed to general anaesthesia and in 14% to regional anaesthesia. In almost all mothers undergoing anaesthesia, propofol and/or sevoflurane with standard ASA monitoring were used. Mothers underwent a broad spectrum of surgery, with abdominal procedures the most frequent type (41.9%) and appendicectomy the most common surgical operation (22.5%).

The primary outcome (t-score of the composite score of the executive function questionnaire, BRIEF) was not significantly different between both groups, with the weighted mean difference (95%CI) of t-scores being 1.9 (−0.4–4.2), p = 0.10 (Fig. 2a). Likewise, groups did not differ for the secondary outcomes from the psychosocial problem questionnaire (CBC). The weighted mean difference (95%CI) of t-scores were: 0.4 (−2.0–2.8), p = 0.74 for internalising problems; 0.4 (−1.7–2.5), p = 0.68 for externalising problems; and 1.1(−1.2–3.3), p = 0.36 for total problems. There were no significant differences in the risks for DSM-5 [5] diagnoses or for learning disorders (Fig. 3).

Several sensitivity analyses and additional analyses were performed. When statistical methods to reduce bias by confounders were used, there were no significant differences, regardless of the mode of analysis. For example, when methods other than inverse probability of treatment weighting were used to reduce bias by confounders, all conclusions persisted (Fig. 2b and c, online Supporting Information, Figures S3a and b). Likewise, groups did not differ when assessing the risk for t-scores ≥ 60 (Fig. 4a–c), after exclusion of children with potential confounders (see online Supporting Information, Figure S4) and when additionally taking into account other potential confounders such as gestational age and birth weight (online Supporting Information, Figure S5). There was also no difference between groups when repeating the analyses separately for the age groups of 2–5 y and 6–18 y (online Supporting Information, Figures S6 and S7) and when only considering raw scores (online Supporting Information, Figures S6 and S7 and Appendix S2).

In the sensitivity analyses, exposure to anaesthesia was only associated with an impaired outcome when confounders were not taken into account: t-scores were significantly higher in the exposed group for the composite score of the executive function questionnaire, where the unweighted mean difference (95%CI) was 3.4 (1.2–5.7, p = 0.003), and in the externalising and total problem domains of the psychosocial problem questionnaire (Fig. 2d). The risk for t-scores ≥ 60 was also significantly

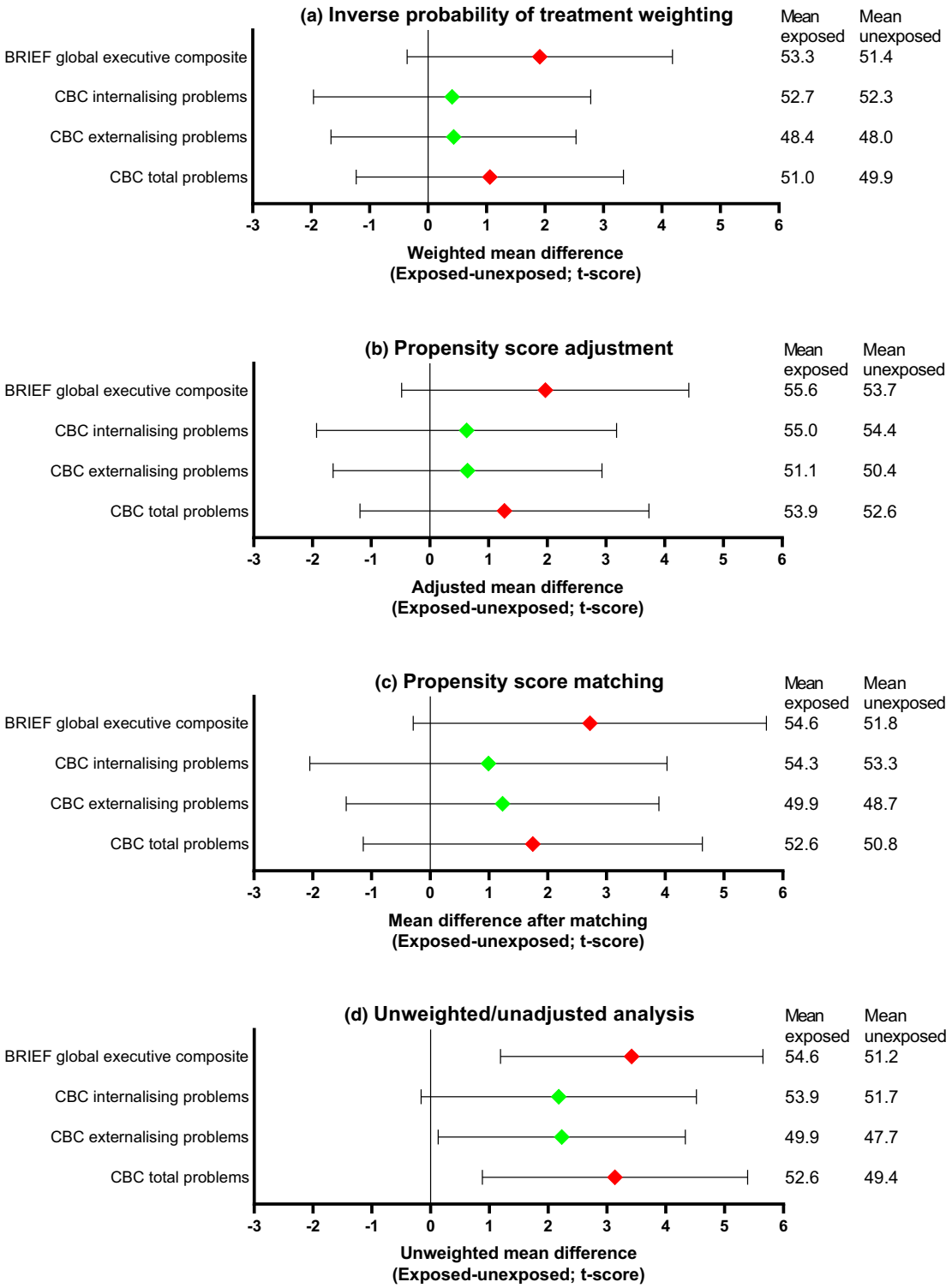

Figure 2 Primary and secondary outcomes. Diamonds and error bars represent the estimate for the mean difference of t-scores (exposed minus unexposed) and their 95%CIs. Inverse probability of treatment weighting (panel a) was used to reduce bias by confounders. In the sensitivity analyses, other methods to reduce bias by confounders were used (panels b and c). Data were also analysed without taking confounders into account (panel d). BRIEF, Behavior Rating Inventory of Executive Function; CBC, Child Behavior Checklist.

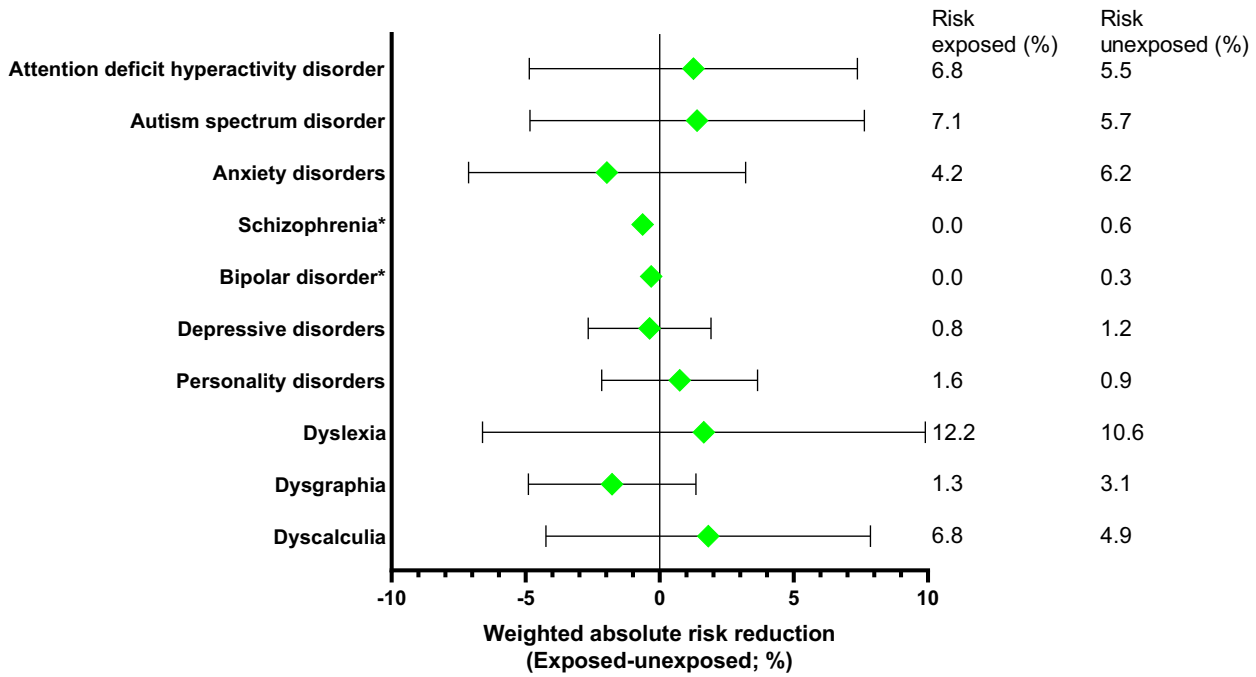

Figure 3 Weighted absolute risk reduction for DSM-5 diagnoses and learning disorders in 85 children exposed to anaesthesia and 313 not exposed. Diamonds and error bars show the estimate for the weighted absolute risk reduction (exposed minus unexposed) and their 95% CIs. *Only the unweighted absolute risk reduction without CI could be calculated, because the risk was zero in the exposed group.

increased for the composite score of the executive function questionnaire (BRIEF), the unweighted absolute risk reduction (95%CI) being 11.7 (0.9–22.5), p = 0.03 and the psychosocial problem questionnaire (CBC) total problems (Fig. 4d). In contrast, groups did not differ for the psychosocial problem questionnaire internalising problems (Fig. 2d) nor for the incidence of DSM-5 [5] diagnoses and learning disorders (see online Supporting Information, Figure S3c).

Several exploratory analyses were performed. For the composite score of the executive function questionnaire, but not for the psychosocial problem questionnaire total problems, significantly higher t-scores were obtained for children in the subgroups exposed to general anaesthesia, intra-abdominal surgery, prolonged anaesthesia (>1 h) and laparoscopic surgery. Estimates for the weighted mean difference of t-scores in these subgroups ranged from 3.2 to 4.5 (online Supporting Information, Figure S8).

The effect size of t-score of the composite score of the executive function questionnaire and the psychosocial problem questionnaire total problems associated with exposure to anaesthesia was of a magnitude comparable with the effect size of parental university level, maternal age at birth and renting vs. owning the family residence (Fig. 5 and online Supporting Information, Figure S9).

Both significantly increased t-scores and a significantly increased risk for t-scores ≥ 60 were observed in children aged 2–5 y for sleeping problems (online Supporting Information, Figure S6) and in children aged 6–18 y for working memory, plan/organise and attention problems (online Supporting Information, Figure S7).

Discussion

We suggest that, in the general population, prenatal exposure to anaesthesia for non-obstetric surgery is not associated with clinically meaningful impairments in neurodevelopmental outcomes. We make this assertion considering the observed effect size and make the following comments to support this. First, the observed effect size was smaller than the standard deviation of t-scores (i.e. 10). Second, the observed effect size was comparable with the effect size of the parents' university level, mother's age and family residence ownership. Third, comparable effect sizes (adjusted mean difference of t-score from the composite of the executive function questionnaire of 2.04) have been reported for children exposed to > 2 h vs. ≤ 2 h of screen time per day [18]. Fourth, overall, we found no evidence for an increased risk of having a t-score ≥ 60. Additionally, there were no significant differences between children exposed and unexposed in utero in any of our secondary outcomes.

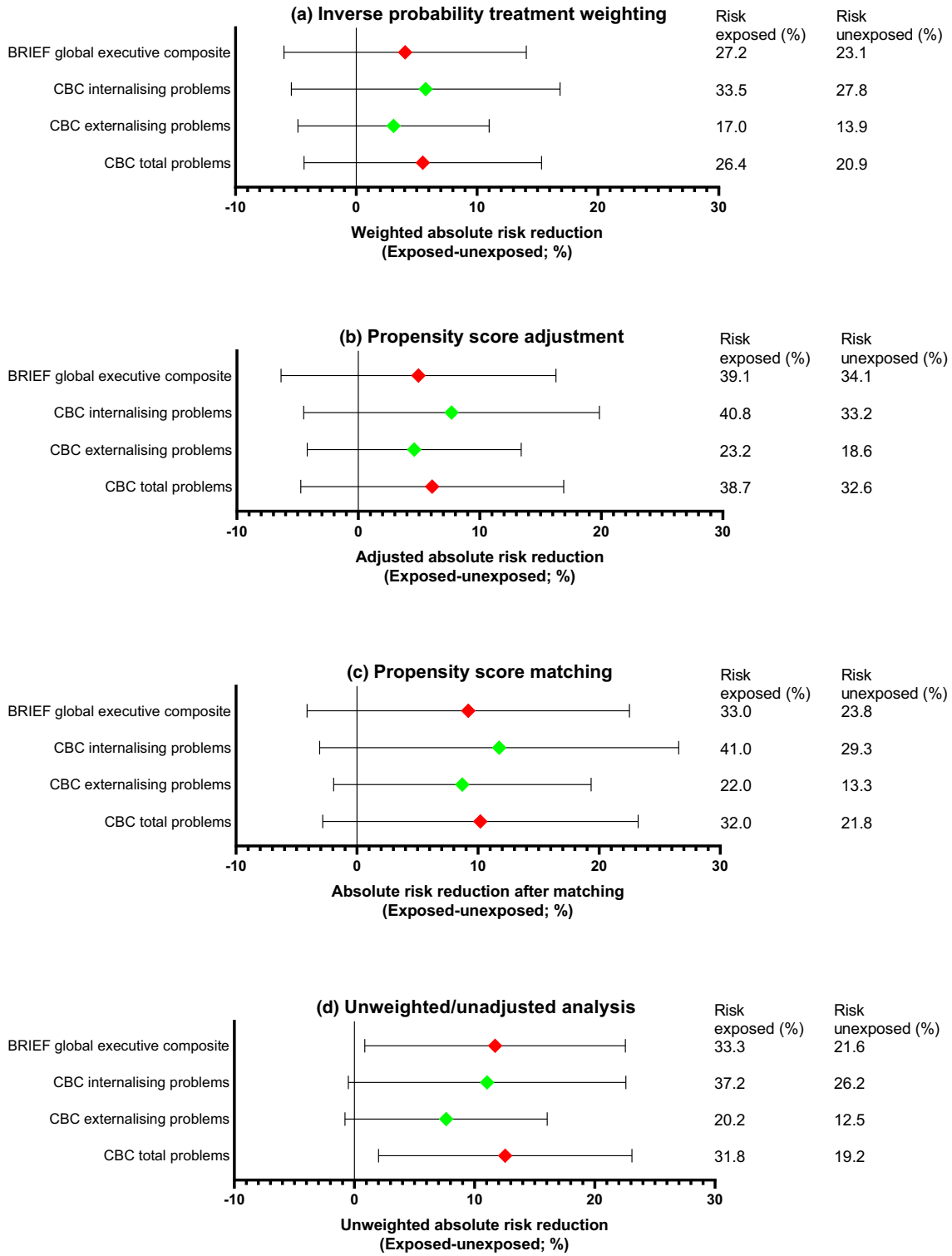

Figure 4 T-scores ≥ 60 on the executive function and psychosocial problem questionnaire (at risk/clinically elevated). Diamonds and error bars show the estimate for the weighted absolute risk reduction (exposed minus unexposed) and their 95% CIs. Inverse probability of treatment weighting (panel a) was used to reduce bias by confounders. In the sensitivity analysis, other methods were used (panel b, c) to take confounders into account. Panel d shows the unweighted/unadjusted analysis. BRIEF, Behavior Rating Inventory of Executive Function; CBC, Child Behavior Checklist.

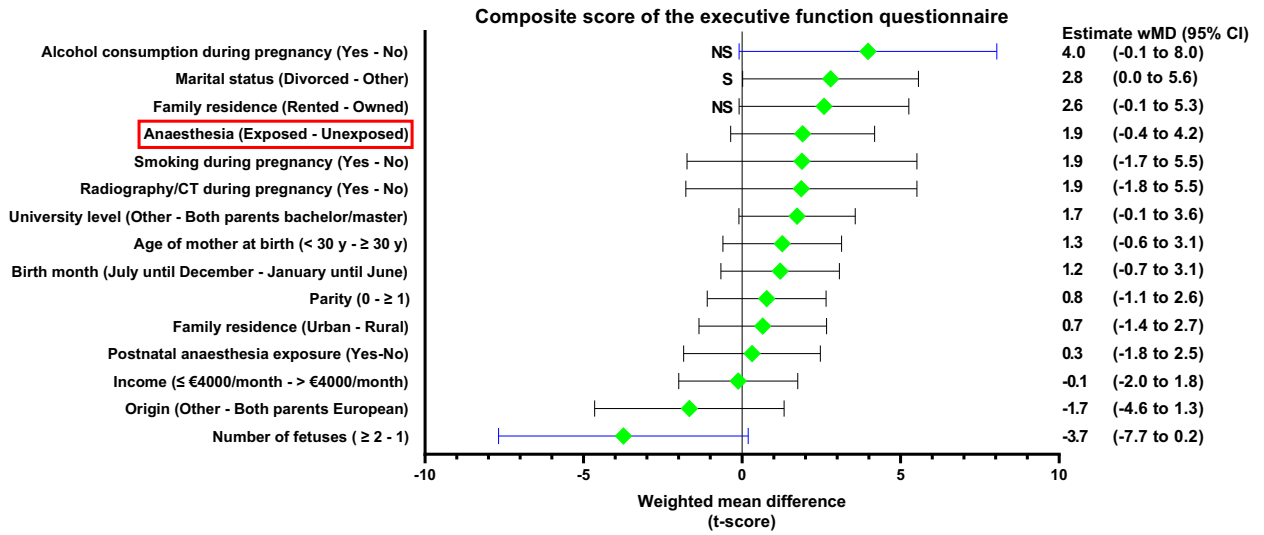

Figure 5 Ranked comparison of the effect size (95%CI) of the primary outcomes with all weighted confounders. NS, not significant; S, significant, wMD, weighted mean difference of t-scores.

A recent study reported that in-utero exposure to general anaesthesia was associated with more externalising psychosocial problems, but not other impairments [17]. However, this study has several limitations, including: an uncertain exposure status; a small sample size (22 likely exposed children); and anaesthesia exposure dating back to an era (1989–1992) that predates modern anaesthetic drugs and monitoring standards. Our study overcomes these limitations: only children with a certain exposure status receiving modern anaesthetics and standard ASA monitoring were included, and the sample size was statistically justified.

A systematic review of studies investigating the effects of anaesthesia exposure after birth concluded that anaesthesia seems to be a marker rather than a cause of impaired neurodevelopment [19]. The effect size for anaesthesia exposure after birth has been reported to be smaller than the effect size caused by gender, birth month of the same year, maternal education and preterm birth [20, 21]. The observed effect sizes for the executive function and psychosocial problem questionnaires in our study were comparable with those in studies prospectively assessing neurodevelopmental outcomes after a single exposure of young children to anaesthesia [22].

Differences between exposed and unexposed children were not limited to the exposure to anaesthesia. Pregnant women receive anaesthesia almost exclusively to enable surgery [1] to treat an underlying condition (e.g. malignant or infectious disease). The underlying pathology will also require diagnostic intervention (e.g. CT scans), other non-surgical treatments (e.g. chemotherapy) or can have other consequences (e.g. social isolation). Some diseases can also be related to socio-economic factors (e.g. low university level). All these factors may affect the neurodevelopmental outcomes of a child. In fact, socio-economic status and other baseline characteristics in the present study were significantly different between the parents of exposed and unexposed children.

To account for the confounders present at baseline, inverse probability of treatment weighting was used to analyse all primary and secondary outcomes to reduce possible bias. Without addressing these confounders by this weighting method, several statistically significant and clinically relevant differences (t-scores ≥ 60) were observed. However, in the subgroups of children whose mothers underwent laparoscopic and intra-abdominal surgery (both probably reflecting more complex pathology and surgery), increased t-scores were observed even after weighting. It is tempting to speculate that these observations may suggest that factors other than anaesthesia are more relevant to neurodevelopmental outcomes.

The present study examined exposures during a broad spectrum of procedures (e.g. abscess incision vs. cardiac surgery), using both general and regional anaesthesia, and the primary outcome was the sum of different cognitive domains. Not observing a difference in the average of this sum does not preclude possibly larger impairments in vulnerable patients, for general anaesthesia or in specific cognitive domains. Larger effect sizes were observed for specific subgroups, and increased risks for t-scores ≥ 60 were detected for specific cognitive subdomains. Also, for the subgroup of general anaesthesia, statistically significant

differences were observed, but the effect size does not represent clinically meaningful impairments. All these differences should be interpreted as hypothesis-generating as the chance of type-1 errors is high. As 90% of children were exposed to a single anaesthetic with an average duration of 91 min, it would be reasonable to state that the primary outcome conclusions refer to this type of exposure and cannot be extrapolated to repeated or longer exposures.

None of the maternal procedures performed in the exposed group could have been postponed until after delivery. For example, delayed treatment of appendicitis could result in a fetal loss rate of 20–35% [23] and maternal septic shock [24]. It is reassuring that the observed effect sizes for this study's primary and secondary outcomes were small and clinically insignificant. The results of our study do not change the recommendation that all urgent and essential procedures should be performed immediately during pregnancy [25].

We acknowledge that our study has several limitations. First, the executive function and the psychosocial problem questionnaires assess parentally reported changes in behaviour and are, therefore, theoretically incapable of detecting smaller neurocognitive impairments not affecting behaviour. Additionally, the answers of the parents may be subjective. However, several studies prospectively measuring the effects of anaesthesia exposure after birth used both the executive function and the psychosocial problem questionnaires and more advanced direct neuropsychological testing (e.g. Wechsler Abbreviated Scale of Intelligence) [22]. In these studies, the effect sizes observed by these more advanced instruments were smaller than the effect sizes detected by the questionnaires [22]. This suggests that the conclusions of the present study would have been probably altered to only a minor extent if more advanced neuropsychiatric testing had been used. Future studies should assess neurodevelopmental outcomes using more advanced neuropsychological testing to confirm this hypothesis. Second, it is not possible to disentangle the effects of anaesthesia from those of surgery, the underlying pathology and other unknown factors. Any impact of anaesthesia could only be investigated by randomising pregnant women to general or regional anaesthesia. Third, the study was performed in a tertiary referral centre, in which high-risk patients may be over-represented when compared with the general population. However, the average t-scores of both groups were close to the average t-score (i.e. 50) of the reference population used in the validation of the executive function and psychosocial problem questionnaires. Fourth, a wide spectrum of procedures was included in this study. Future studies could focus on more homogenous patient groups, e.g. intra-abdominal surgery. Fifth, it was impossible to increase the response rate beyond 64%. However, it has been shown that the response rate and non-response bias are not correlated. Consequently, estimating a possible non-response bias is more important than achieving a higher response rate [11, 12]. In this study, it can be concluded that the non-response bias is negligible.

This study also has some strengths. It is the first study to investigate the effects of in-utero exposure to modern anaesthesia techniques and the sample size was statistically justified. Inverse probability of treatment weighting was used to reduce bias by confounders with multiple sensitivity analyses confirming the robustness of the conclusions. The executive function questionnaire [9, 10, 13], the psychosocial problem questionnaire [9, 10, 13–15, 26], DSM-5 [5] diagnoses and the presence of learning disorders [27–32] are established tools for the assessment of neurodevelopmental outcomes in studies investigating postnatal anaesthesia exposure. The validity of our study design is illustrated to some degree by the fact that clinically small differences could be detected with statistical significance when confounders were not considered. In conclusion, we found no evidence in the general population for an association between prenatal exposure to anaesthesia and impaired neurodevelopmental outcomes.

Acknowledgements

The protocol (including the statistical analysis plan) was approved by the local ethics committee (Commissie Medische Ethiek, University Hospitals Leuven, Belgium). We thank S. Fieuws, M. van Hoef, C. Huygens, E. Michiels, E. Stabel, M. Allegaert, K. Stroobants, A. Brenders, L. Van Eylen, S. Maris, J. De Wulf and C. Ons, and their roles are described in online Supporting Information Appendix S3. TB, MVDV, SD and SR are funded by the European Society of Anaesthesiology and Intensive Care Young Investigator Start-up Grant 2021, the Society for Anaesthesia and Resuscitation of Belgium research grant 2019 and the Obstetric Anaesthetists' Association International Grant 2019. JD is funded by the Great Ormond Street Hospital Charity Fund, the Wellcome Trust and Engineering and Physical Sciences Research Council. No competing interests declared.

References

1. Devroe S, Bleeser T, Van de Velde M, et al. Anesthesia for non-obstetric surgery during pregnancy in a tertiary referral center: a 16-year retrospective, matched case-control, cohort study. *International Journal of Obstetric Anesthesia* 2019; **39**: 74–81.
2. Cho S, Chung RK, Jin SH. Factors affecting maternal and fetal outcomes of non-obstetric surgery and anesthesia during

pregnancy: a retrospective review of data at a single tertiary university hospital. *Journal of Korean Medical Science* 2020; **35**: e113.
3. Bleeser T, Van Der Veeken L, Fieuws S, et al. Effects of general anaesthesia during pregnancy on neurocognitive development of the fetus: a systematic review and meta-analysis. *British Journal of Anaesthesia* 2021; **126**: 1128–40.
4. Food and Drug Administration. FDA drug safety communication: FDA review results in new warnings about using general anesthetics and sedation drugs in young children and pregnant women. 2016. https://www.fda.gov/Drugs/DrugSafety/ucm532356.htm (accessed 11/08/2022).
5. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th edn. Arlington, Washington DC: American Psychiatric Association, 2013.
6. Hennessy S, Bilker WB, Berlin JA, Strom BL. Factors influencing the optimal control-to-case ratio in matched case-control studies. *American Journal of Epidemiology* 1999; **149**: 195–7.
7. Roth RM, Isquith PK, Gioia GA. Assessment of executive functioning using the behavior rating inventory of executive function (BRIEF). In Goldstein S, Naglieri J, eds. *Handbook of Executive Functioning*. New York, NY: Springer, 2014.
8. Achenbach TM, Rescorla LA. *Manual for the ASEBA School-Age Forms and Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families, 2001 https://aseba.org/wp-content/uploads/2019/01/ASEBA-Reliability-and-Validity-School-Age.pdf (accessed 27/09/2022).
9. 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.
10. 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.
11. Halbesleben JR, Whitman MV. Evaluating survey quality in health services research: a decision framework for assessing nonresponse bias. *Health Services Research* 2013; **48**: 913–30.
12. Phillips AW, Reddy S, Durning SJ. Improving response rates and evaluating nonresponse bias in surveys: AMEE guide No. 102. *Medical Teacher* 2016; **38**: 217–28.
13. 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. *Journal of the American Medical Association* 2016; **315**: 2312–20.
14. Bakri MH, Ismail EA, Ali MS, Elsedfy GO, Sayed TA, Ibrahim A. Behavioral and emotional effects of repeated general anesthesia in young children. *Saudi Journal of Anaesthesia* 2015; **9**: 161–6.
15. Ing C, DiMaggio C, Whitehouse A, et al. Long-term differences in language and cognitive function after childhood exposure to anesthesia. *Pediatrics* 2012; **130**: e476–485.
16. Walkden GJ, Pickering AE, Gill H. Assessing long-term neurodevelopmental outcome following general anesthesia in early childhood: challenges and opportunities. *Anesthesia and Analgesia* 2019; **128**: 681–94.
17. Ing C, Landau R, DeStephano D, Miles CH, von Ungern-Sternberg BS, Li G, Whitehouse AJO. Prenatal exposure to general anesthesia and childhood behavioral deficit. *Anesthesia and Analgesia* 2021; **133**: 595–605.
18. Vohr BR, McGowan EC, Bann C, Das A, Higgins R, Hintz S. Association of high screen-time use with school-age cognitive, executive function, and behavior outcomes in extremely preterm children. *Journal of the American Medical Association Pediatrics* 2021; **175**: 1025–34.
19. Foubert R, Devroe S, Foubert L, Van de Velde M, Rex S. Anesthetic neurotoxicity in the pediatric population: a systematic review of the clinical evidence. *Acta Anaesthesiologica Belgica* 2020; **71**: 51–65.
20. de Heer IJ, Tiemeier H, Hoeks SE, Weber F. Intelligence quotient scores at the age of 6 years in children anaesthetised before the age of 5 years. *Anaesthesia* 2017; **72**: 57–62.
21. Glatz P, Sandin RH, Pedersen NL, Bonamy AK, Eriksson LI, Granath F. Association of anesthesia and surgery during childhood with long-term academic performance. *Journal of the American Medical Association Pediatrics* 2017; **171**: e163470.
22. Ing C, Jackson WM, Zaccariello MJ, et al. Prospectively assessed neurodevelopmental outcomes in studies of anaesthetic neurotoxicity in children: a systematic review and meta-analysis. *British Journal of Anaesthesia* 2021; **126**: 433–44.
23. Palanivelu C, Rangarajan M, Senthilkumaran S, Parthasarathi R. Safety and efficacy of laparoscopic surgery in pregnancy: experience of a single institution. *Journal of Laparoendoscopic and Advanced Surgical Techniques* 2007; **17**: 186–90.
24. Abbasi N, Patenaude V, Abenhaim HA. Management and outcomes of acute appendicitis in pregnancy-population-based study of over 7000 cases. *British Journal of Obstetrics and Gynaecology* 2014; **121**: 1509–14.
25. Auron M, Duran Castillo MY, Garcia OFD. Perioperative management of pregnant women undergoing nonobstetric surgery. *Cleveland Clinic Journal of Medicine* 2020; **88**: 27–34.
26. 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.
27. DiMaggio C, Sun LS, Li G. Early childhood exposure to anesthesia and risk of developmental and behavioral disorders in a sibling birth cohort. *Anesthesia and Analgesia* 2011; **113**: 1143–51.
28. Bong CL, Allen JC, Kim JT. The effects of exposure to general anesthesia in infancy on academic performance at age 12. *Anesthesia and Analgesia* 2013; **117**: 1419–28.
29. Ko WR, Liaw YP, Huang JY, et al. Exposure to general anesthesia in early life and the risk of attention deficit/hyperactivity disorder development: a nationwide, retrospective matched-cohort study. *Pediatric Anesthesia* 2014; **24**: 741–8.
30. Ko WR, Huang JY, Chiang YC, et al. Risk of autistic disorder after exposure to general anaesthesia and surgery: a nationwide, retrospective matched cohort study. *European Journal of Anaesthesiology* 2015; **32**: 303–10.
31. Nestor KA, Zeidan M, Boncore E, Richardson A, Alex G, Weiss M, Islam S. Neurodevelopmental outcomes in infants undergoing general anesthesia. *Journal of Pediatric Surgery* 2017; **52**: 895–900.
32. Ing C, Sun M, Olfson M, DiMaggio CJ, Sun LS, Wall MM, Li G. Age at exposure to surgery and anesthesia in children and association with mental disorder diagnosis. *Anesthesia and Analgesia* 2017; **125**: 1988–98.

Supporting Information

Additional supporting information may be found online via the journal website.

Appendix S1. Further description of methods.
Appendix S2. Further description of results.
Appendix S3. Acknowledgements.
Figure S1. Contact with parents and response rate.
Figure S2. Inverse probability of treatment weighting.

Figure S3. Diagnostic and Statistical Manual of Mental Disorders diagnoses and learning disorders.

Figure S4. Result after exclusion of children for which potential confounders could bias the results.

Figure S5. Results after additionally using other variables to calculate propensity scores.

Figure S6. The preschool executive function questionnaire and the preschool psychosocial problem questionnaire: children aged 2–5 y.

Figure S7. The school age executive function questionnaire and the school age psychosocial problem questionnaire: children aged 6–18 y.

Figure S8. Subgroup analyses.

Figure S9. Effect sizes for the psychosocial problem questionnaire total problems.

Table S1. Comparison of baseline characteristics of respondents with those of non-respondents.

Table S2. Baseline characteristics of study participants.