Neurodevelopmental outcome at 5 years of age after general anaesthesia or awake-regional anaesthesia in infancy (GAS)

McCann, Mary Ellen; de Graaff, Jurgen C; Dorris, Liam; Disma, Nicola; Withington, Davinia E; Bell, Graham; Grobler , Anneke; Stargatt, Robyn; Hunt, Rodney W; Sheppard , Suzette J. ; Marmor, Jacki; Giribaldi, Gaia; Bellinger, David C; Hartmann, Penelope L; Hardy, Pollyanna; Frawley, Geoff; Izzo , Francesca ; von Ungern Sternberg, Britta S; Lynn, Anne; Wilton, Niall

DOI:
10.1016/S0140-6736(18)32485-1

Document Version
Peer reviewed version

Citation for published version (Harvard):
McCann, ME, de Graaff, JC, Dorris, L, Disma, N, Withington, DE, Bell, G, Grobler , A, Stargatt, R, Hunt, RW, Sheppard , SJ, Marmor, J, Giribaldi, G, Bellinger, DC, Hartmann, PL, Hardy, P, Frawley, G, Izzo , F, von Ungern Sternberg, BS, Lynn, A, Wilton, N, Mueller, M, Polaner, D, Absalom, AR, Szmuk, P, Morton, NS, Berde, C, Soriano, S & Davidson, AJ 2019, 'Neurodevelopmental outcome at 5 years of age after general anaesthesia or awake-regional anaesthesia in infancy (GAS): an international, multicentre, randomised, controlled equivalence trial', *The Lancet*, vol. 393, no. 10172, pp. 664-677. https://doi.org/10.1016/S0140-6736(18)32485-1

Neurodevelopmental outcome at 5 years of age after general anaesthesia or awake-regional anaesthesia in infancy (GAS): an international, multicentre, randomised controlled equivalence trial.

Authors: and The GAS Consortium

1. Mary Ellen McCann MD 1

2. Jurgen C. de Graaff PhD 2, 3

3. Liam Dorris DClinPsy 4, 5

4. Nicola Disma MD 6

5. Davinia Withington BM 7, 8 *

6. Graham Bell MBChB 9

7. Anneke Grobler PhD 10, 11

8. Robyn Stargatt PhD 12, 13

9. Rodney W. Hunt PhD 11, 14, 15 *

10. Suzette J. Sheppard BSc 16

11. Jacki Marmor MEd 17

12. Gaia Giribaldi MD 6

13. David C Bellinger PhD 17 *

14. Penelope L Hartmann PhD 16, 18

15. Pollyanna Hardy MSc 19

16. Geoff Frawley MBBS 11, 16, 20

17. Francesca Izzo MD 21

18. Britta S von Ungern Sternberg PhD 22, 23, 24 *

19. Anne Lynn MD 25, 26 *

20. Niall Wilton MBBS 27

21. Martin Mueller MD 28

22. David M. Polaner MD 29, 30 *

23. Anthony R. Absalom MBChB 31 *

24. Peter Szmuk MD 32, 33 *

25. Neil Morton MD 9, 34

26. Charles Berde MD 1 *

27. Sulpicio Soriano MD 1 *

28. Andrew J. Davidson MD 11, 16, 20 *

Institutions

1. Department of Anesthesiology, Critical Care and Pain Medicine, Boston Children's Hospital, Boston, Massachusetts, US

2. Department of Anaesthesiology, Erasmus Medical Center, Rotterdam, The Netherlands

3. Department of Anaesthesiology, University Medical Centre Utrecht, Utrecht, The Netherlands

4. Paediatric Neurosciences, Royal Hospital for Children, Glasgow, Scotland, UK.

5. Institute of Health & Wellbeing, University of Glasgow, Glasgow, Scotland, UK.

6. Department of Anaesthesia, Istituto Giannina Gaslini, Genoa, Italy

7. Department of Anaesthesia, Montreal Children's Hospital, Montreal, Canada

8. Department of Anesthesia, McGill University, Montreal, Canada

9. Department of Anaesthesia, Royal Hospital for Children, Glasgow, Scotland, UK

10. Clinical Epidemiology and Biostatistics Unit, Murdoch Children's Research Institute, Melbourne, Australia.

11. Department of Paediatrics, University of Melbourne, Melbourne, Australia

12. School of Psychological Science, La Trobe University, Victoria, Australia

13. Child Neuropsychology, Murdoch Children's Research Institute, Melbourne, Australia

14. Department of Neonatal Medicine, The Royal Children's Hospital, Melbourne Australia

15. Neonatal Research Group, Murdoch Children's Research Institute, Melbourne, Australia

16. Anaesthesia and Pain Management Research Group, Murdoch Children's Research Institute, Melbourne, Australia

17. Department of Neurology, Boston Children's Hospital, Boston, Massachusetts, US

18. School of Behavioural & Health Sciences, Australian Catholic University, Melbourne, Australia

19. Birmingham Clinical Trials Unit, University of Birmingham, UK

20. Department of Anaesthesia and Pain Management, The Royal Children's Hospital, Melbourne, Australia

21. Department of Anaesthesiology and Paediatric Intensive Care, Ospedale Pediatrico Vittore Buzzi, Milan, Italy

22. Medical School, The University of Western Australia, Perth, Australia

23. Department of Anaesthesia and Pain Management, Perth Children's Hospital, Perth, Australia

24. Telethon Kid's Institute, Perth, Australia

25. Department of Anesthesiology and Pain Medicine, and Pediatrics University of Washington, Seattle, Washington, US

26. Department of Anaesthesia and Pain Medicine, Seattle Children's Hospital, Seattle, Washington, US

27. Department of Paediatric Anaesthesia and Operating Rooms, Starship Children's Hospital, Auckland District Health Board, Auckland, New Zealand

28. Department of Anaesthesia, The University of Iowa Hospitals and Clinics, Iowa City, Iowa, US

29. Department of Anaesthesiology, Children's Hospital Colorado, Denver, Colorado, US

30. Department of Anaesthesiology, University of Colorado, Denver, Colorado, US

31. Department of Anaesthesiology, University Medical Centre Groningen, Groningen University, Groningen, The Netherlands

32. Department of Anesthesiology and Pain Management, University of Texas Southwestern and Children's Medical Centre Dallas, Dallas, Texas, US

33. Department of Outcomes Research, Cleveland Clinic, Cleveland, Ohio, US

34. Academic Unit of Anaesthesia, Pain and Critical Care, University of Glasgow, Glasgow, UK

Corresponding Author:

Andrew J. Davidson, Anaesthesia and Pain Management Research Group, Murdoch Childrens Research Institute, The Royal Children's Hospital, Flemington Road, Parkville, Victoria, 3052. Australia. Email Address: andrew.davidson@rch.org.au Phone Number: +61 (0)3 9345 4008

The corresponding author attests that all authors have seen and approved the final text.

Summary: (Word count: 365)

Background: In laboratory animals, exposure to most general anaesthetics leads to neurotoxicity manifested by neuronal cell death, and abnormal behaviour and cognition. Some, ~~everal~~ large human cohort studies demonstrate an association between general anaesthesia at a young age and subsequent neurodevelopmental deficits, but are prone to bias. Others have found no evidence for an association. We aimed to establish whether general anaesthesia in early infancy has an effect on neurodevelopmental outcomes in a randomised controlled trial (RCT).

Methods: In this international assessor-masked equivalence RCT, infants less than 60 weeks' postmenstrual age and born at greater than 26 weeks gestation undergoing inguinal herniorraphies without prior exposure to general anaesthesia or risk factors for neurologic injury were recruited. They were randomly assigned to receive either an awake-regional or sevoflurane-based general anaesthetic. The primary outcome measure was the Wechsler Preschool and Primary Scale of Intelligence-Third Edition (WPPSI-III) Full Scale Intelligence Quotient (FSIQ) at 5 years of age. The primary analysis was as-per-protocol adjusted for gestational age at birth and country using multiple imputation to deal with missing data. An intention-to-treat analysis was also performed. A difference in means of five points was predefined as the clinical equivalence margin. This trial is registered with ANZCTR, number ACTRN12606000441516 and ClinicalTrials.gov, number NCT007566000.

Findings: Between Feb 2007 and Jan 2013, 722 infants were randomised, 363 to the awake-regional and 359 to general anaesthesia. The median duration of anaesthesia in the general anaesthetic group was 54 minutes. There were 74 protocol violations in the awake-regional group and 2 in the general anaesthesia group. Primary outcome data for the as-per-protocol analysis were obtained from 205 children in the awake-regional group and 242 in the general anaesthesia group. The FSIQ score (mean [standard deviation (SD)]) was 99.08 (18.35) in the awake-regional group and 98.97 (19.66) in the general anaesthesia group, with a difference in means (awake-regional minus general anaesthesia) of 0.23, 95% Confidence Intervals -2.59 to 3.06) showing strong evidence of equivalence. The results with the intention-to-treat analysis were similar to the as-per-protocol analysis.

Interpretation: We found strong evidence that just under an hour of general anaesthesia in early infancy does not alter neurodevelopmental outcome compared to awake-regional anaesthesia in a predominantly male study population.

Funding: National Institutes of Health (NIH) USA, Food and Drug Administration USA, Thrasher Research Fund, Australia National Health and Medical Research Council (NHMRC), Health Technologies Assessment-National Institute for Health Research UK. Australian and New Zealand College of Anaesthetists, Murdoch Children's Research Institute, Canadian Institutes of Health Research, Canadian Anesthesiologists Society, Pfizer Canada, Italian Ministry of Health (RF-2011-02347532), Fonds NutsOhra, the UK Clinical Research Network (UKCRN) and departmental sources. Britta S von Ungern-Sternberg is partly funded by the Perth Children's Hospital Foundation, the Stan Perron Charitable Trust, and the Callahan Estate.

Research in context

Evidence before this study

We searched Medline and Cochrane controlled trials register (May 20, 2018) for original research and meta-analyses describing the association between anaesthetic exposure during childhood and neurodevelopmental outcome.

The search terms used were "anesthesia" and "child development" or "anesthesia" and "learning disorders". No randomised trials were found except for the interim analysis of this trial published in the Lancet in 2016 which found equivalence in Bayley-III scores between infants exposed to either regional or general anaesthesia. The majority of large cohort studies report an association between surgery before the age of four years and an increased risk for a later diagnosis of a behavioural problem or poorer academic attainment. In some of the studies the size of the increased risk is very small, in others it is only seen after multiple exposures. Several, but not all, of the cohort studies did not find an association with neurocognitive outcome as assessed by formal IQ testing. Weaknesses in these cohort studies include confounding, bias, heterogeneous populations at the time of exposure and heterogeneous outcome measures making interpretation and generalisation problematic.

Added value of this study

We report the 5 year neurodevelopmental outcome results for the GAS trial, the first randomised controlled trial designed to assess the effect of general anaesthesia in infancy on neurodevelopmental outcome. We used the most reliable and validated measure of general intellectual ability, the Wechsler Preschool and Primary Scale of Intelligence-Third Edition Full Scale IQ score and *found strong evidence for equivalence between awake-regional and just less than one hour of general anaesthesia*. No significant differences were seen in a range of other neurocognitive and behavioural measures.

Implications of all the available evidence

This randomised controlled trial provides strong evidence that an hour of exposure to a general anaesthetic during early infancy does not cause measureable neurocognitive or behavioural deficits at 5 years of age. These results are consistent with the MASK and PANDA cohort

studies. Nearly half the general anaesthetics in infancy are under an hour in duration and thus this study should allay some of the concerns generated by the preclinical data and previous cohort studies. This trial does not address the possibility that longer or repeated anaesthesia exposures in early childhood are detrimental. The trial was also conducted in a predominantly male population, and thus further research is needed which is directed specifically towards answering these questions relating to female sex, and multiple and prologed expsoures.

Introduction

There are ongoing concerns about anaesthesia induced neurotoxicity for the developing brain.[1-2] In animal models, exposure to most general anaesthetics at a young age results in a range of morphologic changes.[3] These exposed animals, including non-human primates, exhibit neuronal cell death, impaired neurogenesis, glial death and abnormal axon formation.[4-7] Some animal models have also found that anaesthesia exposure in infancy is associated with altered behaviours including heightened emotional reactivity to threats, and impaired learning and memory formation persisting into early adulthood.[8,9] Given the greater complexity of human development, it is unclear how these animal model findings translate to humans.

In human cohort studies there is mixed and conflicting evidence for an association between exposure to anaesthesia in early childhood and a range of adverse neurodevelopmental outcomes.[10] In light of the preclinical and clinical findings~~, anaesthesia societies in several countries have issued statements advising practitioners to consider delaying non-urgent surgery and to be prepared to discuss the issue with parents and~~ the United States Food and Drug Administration has mandated warning labels on most general anaesthetics used in children. [11,12] There have also been numerous calls for more definitive research to determine if anaesthetic exposure in early childhood has a clinically relevant impact on neurodevelopment in humans.[13,14]

There are inherent difficulties in drawing any conclusions about causation from these cohort studies due to likely confounding, hence a randomised controlled trial would provide the strongest evidence for or against general anaesthesia causing adverse neurodevelopmental outcome.

The neurodevelopmental outcome after general anaesthesia or awake-regional anaesthesia in infancy (GAS) trial was designed to answer the question of whether an exposure to general anaesthesia ~~exposure~~ in infants leads to clinically significant long term neurodevelopmental changes. A randomised trial to answer this question could only be performed on children undergoing a surgery for which either a volatile anaesthetic (which has been shown to cause injury and neurobehavioural deficits in animal models) or an awake-regional technique (which does not cause neuronal injury in animal models) can be used.[15] Inguinal herniorraphy is one such surgery. An equivalence design was chosen as the primary aim was to determine if we

could exclude general anaesthesia causing clinically relevant neurotoxicity. Our hypothesis was that there would be no clinically important differences in neurodevelopmental outcome between general anesthesia and regional anesthesia. Such a finding of equivalence would result in: a) clinicians no longer subjecting children to the various risks of delaying surgery, and b) anaesthetists not avoiding general anaesthesia by using alternative, and potentially less well established anaesthetic techniques.

The primary outcome for this trial (reported in this paper) is the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III) Full Scale Intelligent Quotient (FSIQ) measured at 5 years of age. A range of other secondary neurodevelopmental outcomes were also assessed at 5 years of age and are reported in this paper. Neurodevelopmental outcome at 2 years of age for the GAS trial was assessed using the Bayley Scales of Infant and Toddler Development III and has been previously published.[16] There was no evidence for a difference in the scores between awake-regional and general anaesthesia groups. An assessment at two years was regarded as an interim or secondary outcome as neurodevelopmental delays can be measured more accurately by assessments conducted at five years of age. Data relating to apnoea in the immediate post-operative period, intra-operative blood pressure, regional anaesthesia and surgical outcomes have been published previously. [17-20]

Methods

Study design

This was a multicentre, international, parallel group, randomised, assessor masked, controlled equivalence trial comparing neurodevelopmental outcome at 5 years of age after infants were randomised to receive awake-regional anaesthesia or general anaesthesia for inguinal herniorraphy. The trial was done in 28 hospitals in Australia, Italy, the US, the UK, Canada, the Netherlands and New Zealand. Institutional Review Board or Human Research Ethics Committee approval was obtained at each site and written informed consent was obtained from the infant's parents or guardians. A summary of the protocol is available online.[21]

The GAS trial is registered in Australia and New Zealand at ANZCTR: ID# ACTRN12606000441516 first registered on 16th October 2006; in the United States (US) at ClinicalTrials.gov: ID#: NCT00756600 first registered on 18th September 2008; and in the United Kingdom (UK) at UK Clinical Research Network (UKCRN) ID#: 6635 (ISRCTN ID#: 12437565; MREC No: 07/S0709/20).

The statistical analysis plan is available at ANZCTR (ID# ACTRN12606000441516). http://www.anzctr.org.au/AnzctrAttachments/1422-GAS%20SAP%205%20years.pdf.

Participants

Inclusion criteria were infants up to 60 weeks' postmenstrual age, born at greater than 26 weeks' gestation and scheduled for inguinal herniorraphy. Exclusion criteria were any contraindication for either anaesthetic technique, a history of congenital heart disease requiring surgery or pharmacotherapy, mechanical ventilation immediately before surgery, known chromosomal abnormalities or other known acquired or congenital abnormalities that might affect neurodevelopment, previous exposure to volatile general anaesthesia or benzodiazepines as a neonate or in the third trimester in utero, any known neurological injury such as cystic periventricular leukomalacia or grade three or four intraventricular haemorrhage, any social or geographical factor that might make follow-up difficult or having a primary language at home in a region where neurodevelopmental tests were not available in that language. We identified eligible infants from operating room schedules or at preadmission clinics and recruited in the clinic or in the preadmission areas of the operating floor.

Randomisation and Masking

Infants were randomly assigned (1:1) to receive either general anaesthesia or awake-regional anaesthesia using a 24 hour web-based randomisation service managed by the Data Management and Analysis Centre, Department of Public Health, University of Adelaide, Australia. Randomisation was done in blocks of two or four in a computer generated random allocation sequence and stratified by site and gestational age at birth: 26-29 weeks and 6 days, 30-36 weeks and 6 days and greater than 37 weeks. The anaesthetist was aware of group allocation but individuals who administered neurodevelopmental assessments were not. Parents who asked about their infant's group allocation were informed and told to mask this information from assessors. After assessments were completed, parents and assessors were asked if they were aware of group allocation.

Procedures

The awake-regional group received a spinal, caudal or combined caudal/spinal anaesthetic according to institutional preferences. Bupivacaine or levobupivacaine at a dose of 0.75 -1mg/kg was administered for spinal anaesthesia. Caudal anaesthesia was with 0.25% bupivacaine or levobupivacaine up to a total dose of 2.5 mg/kg. Several patients in the US in whom it was known that the surgery would take longer than one hour were also administered 3% chloroprocaine via a caudal catheter (loading bolus of 3% chloroprocaine 1 ml/kg over several minutes and then an infusion at 1-2 ml/kg/hr). Additional ilioinguinal and field blocks were performed according to surgical preference. Oral sucrose was given if the child was unsettled but no other pharmacological sedation was permitted. Infants who demonstrated agitation that was not resolved by oral sucrose or in whom the awake-regional anaesthetic was inadequate were treated with sevoflurane. The administration of sevoflurane, nitrous oxide or any other general anaesthetic in this group was considered a protocol violation.

The general anaesthesia group received sevoflurane for induction and maintenance in a mix of air and oxygen. The concentration of sevoflurane, choice of airway device, ventilation technique and use of neuromuscular blocking agents were left to the preference of the anaesthetist. Supplemental opioids and nitrous oxide were not allowed but caudal, ilioinguinal-iliohypogastric or field block with bupivacaine were permitted to provide postoperative analgesia.

Both groups could also be given oral, rectal or intravenous paracetamol. Monitoring and recording were identical in both groups with heart rate, blood pressure, oxygen saturation, and expired sevoflurane concentrations (where applicable) every 5 minutes. In both groups intraoperative serum glucose values were measured after induction; rescue protocols for hypoglycaemia, hypotension and hypoxaemia were applied as appropriate.

Outcome assessments

Neuropsychological assessments were to be undertaken within 4 months of the child turning 5 years of age. The total assessment time was estimated to take approximately 3 hours to complete and assessments were performed at each site by a child psychologist certified to conduct the tests. Quality control was maintained by a national coordinating psychologist. The primary outcome measure was that the Wechsler Preschool and Primary Scale of Intelligence-Third Edition Full Scale Intelligence Quotient WPPSI-III FSIQ score. Other ~~The~~ secondary outcome measures ~~tests~~ used were ~~the Wechsler Preschool and Primary Scale of Intelligence-Third Edition (WPPSI-III),~~ selected NEPSY-II subtests to assess attention and executive function, the Wechsler Individual Achievement Test Second Edition (WIAT-II) or the BVN (the Italian equivalent of the WIAT-II), selected subtests of the Children's Memory Scale (CMS), the Global Executive Composite (GEC) of the Behavior Rating of Executive Function – Preschool Version (BRIEF-P), the Adaptive Behavioral Assessment System Second Edition (ABAS-II) and the Child Behaviour Checklist Caregiver Questionnaire (CBCL). Participatory tests were administered by the psychologist and a parent/caregiver completed the informant report questionnaires. Parents were asked if their child had been diagnosed with cerebral palsy (CP), Autism Spectrum Disorder (ASD), Attention Deficit Hyperactivity Disorder (ADHD), or had any other neurodevelopmental issues. They were also asked if the child had received any neurodevelopmental interventions. Hearing or vision problems were also noted. Demographic data, family structure and medical history since randomisation were recorded, and a brief physical and neurologic examination was done for each patient. All these outcome measures were listed *a prior* in the protocol.

All study data were sent to the Murdoch Children's Research Institute in Melbourne, Australia. All data forms were checked by a research assistant not involved in primary data collection or

entry. Data on test forms that were not completed according to test manual instructions were rejected.

An independent Data Safety Monitoring Committee met approximately every 6 months during recruitment. Site visits were performed by the national coordinating teams for each country annually or biennially, and site visits at the national coordinating sites were done by principal investigators from other nations to check the validity of data. Summary data by allocation were presented to this committee.

Statistical Analysis

The study hypothesis was that the primary outcome, WPPSI-III FSIQ score at 5 years of age, is equivalent in infants who are anaesthetised for inguinal herniorraphy using awake-regional anaesthesia or general anaesthesia. Because this was an equivalence study, the outcome was analysed on an as-per-protocol basis to ensure a conservative estimate of the treatment effect in the direction of non-equivalence. In general it is best practice to analyse outcomes on an intention-to-treat (ITT) basis where all participants are included according to their randomised allocation and issues of selection bias are avoided. In this study there were unavoidable protocol violations, the majority of which were in babies allocated to regional anaesthesia who had some exposure to general anaesthesia particularly if the awake-regional anaesthesia failed. If all infants were analysed according to their randomised allocation in an ITT analysis, this switching from one randomised treatment to the other could dilute the potential effect of general anaesthesia and thus bias the trial towards equivalence.[22]

Equivalence was defined *a priori* as the 95% confidence interval (CI) of the difference in means of the FSIQ lying within minus five and plus five IQ points. Intention-to-treat analyses were also planned. All confidence intervals are two-sided

The sample size was based on the primary outcome; the 5-year follow-up WPPSI-III FSIQ score. Assuming an expected difference of one standardised score point, a standard deviation of 15, and a 90% chance that a 95% CI will exclude a difference of more than five points (the largest difference acceptable to show equivalence), the trial would need 598 infants. The sample size formula used was based on approximations to the normal distribution, and used a two one-sided

test (TOST) procedure. Enrolling roughly 720 participants would allow for 10% loss to follow-up and 10% with a major protocol violation.

We used multiple imputation under a multivariate normal distribution to impute missing outcome data in the primary analysis of all outcomes, with a sensitivity analysis on complete cases only. The mi impute mvn statement in Stata was used to do the multiple imputations. The variables used in the multiple imputation models included baseline, post-randomisation, 2 year cognitive variables and 5 year outcome variables. The following prespecified variables were used as possible predictor variables within the imputation approach (since most of these variables also have missingness, they were also imputed where necessary): Baseline: anaesthesia group, country, sex, gestational age at birth, birth weight, mother received antenatal steroids, mother's education, maternal age < 21; Surgery: need for fluid bolus for hypotension, duration of surgery, significant postoperative apnoea, age at surgery; 2 years: composite cognitive, language, motor and social-emotional score of the Bayley scales of Infant and Toddler Development-Third Edition, any additional anaesthetic exposures since the inguinal herniorraphy, any interventions for neurodevelopmental problems, any other neurological abnormality; 5 years: WPPSI-III FSIQ, any chronic illness, any additional anaesthetic exposures since the inguinal herniorraphy, total length of any readmission to hospital, cerebral palsy, any interventions for neurodevelopmental problems, any other neurological abnormality. With many missing observations these multiple imputation models did not always converge, in which case applicable variables were not included in order to ensure convergence of models. The variables used in the analysis model were always included in the imputation models.

For all continuous outcomes, linear regression was used with the factor variables anaesthesia arm (factor levels: awake regional and general anaesthesia), gestational age at birth and country as fixed effects. Adjusted mean differences are presented with 95% Cis.

All binary outcomes were analysed using generalised linear models (GLM) with binomial link function in order to enable estimation of risk ratios, adjusting for the same factors as for the linear regression. Risk ratios are presented with 95% CIs.

All following subgroup analyses were pre-specified in the statistical analysis plan: country, duration of surgery greater or less than 120 minutes, and age at surgery (greater or less than 70 days). A subgroup analysis by ex-term versus ex-preterm (born at <37 weeks gestation) was also

performed post hoc. P-values for the interactions are presented along with subgroup treatment effect estimates and 95% CIs. All analyses were carried out in Stata (version 14.2).

Role of the funders

The funders of this study had no role in data collection, analysis, interpretation, writing this manuscript or the decision to submit this manuscript. AG has complete access to the data. All other authors have access to the data on request. All authors were responsible for the decision to submit this manuscript.

Results

Between 9th February 2007 and 31st January 2013, 722 infants were recruited and randomised at 28 centres in 7 countries (table 1). There were two misrandomisations and one family withdrew consent after randomisation and before surgery. This left 361 children in the intention-to-treat analysis for the awake-regional group and 358 children in the general anaesthesia group [Figure 1]. Table 2 summarises baseline data for each group and Table 3 summarises demographic data at the 5 year assessment. There were 74 protocol violations in the awake-regional group (the surgeries for 5 children were cancelled and 69 children received some sevoflurane or other general anaesthetic agent) and two protocol violations in the general anaesthesia group (surgery cancelled). The only adverse events during the anaesthesia were related to respiratory complications. These have been previously described in full in a separate publication. [17] There were no other adverse events in either group. The frequency of hypotension has also been described elsewhere. [18]

The 5 year follow up assessments were conducted from 13th March 2012 to 27th April 2018. In total 91 families were lost to follow up in the awake-regional group and 97 in the general anaesthesia group; a follow up rate of 74%. Of those that attended for assessment the WPPSI-III FSIQ was complete for 205 in the awake-regional group and 242 in the general anaesthesia group. Numbers lost to follow up and numbers of complete case assessments are listed for each sit in table 1.

Table 4 summarises the results for the individually administered tests for each group and the differences in means between groups. There was strong evidence for equivalence of the WPSSI-III FSIQ means between awake-regional and general anaesthesia groups in both the as-per-protocol and intention-to-treat analyses using multiple imputation to account for missing data (adjusted mean difference for awake-regional minus general anaesthesia 0.23, 95% CI -2.59 to 3.06 for as-per-protocol analysis; and 0.16, -2.45 to 2.78 for intention-to-treat analyses). There was also evidence for equivalence in the complete cases analyses (adjusted mean difference for awake-regional minus general anaesthesia 0.628, 95% CI -2.093 to 3.349 for as-per-protocol analysis; and 0.266, -2.268 to 2.799 for intention-to-treat analyses). In all these analyses the upper and lower bounds of the 95% confidence intervals were well within the prespecified 5 point equivalence margin. There was also evidence for equivalence of the verbal, performance

and processing speed composite scores of the WPPSI-III, with the 95% confidence intervals around the differences in means again within 5 points in as-per-protocol, intention-to-treat, multiple imputation and complete case analyses. For all the other individually administered secondary outcomes (Table 4) and parent or caregiver reported outcomes (Table 5) none of the 95% confidence intervals around the differences in means were either entirely above or below zero in any of the analyses. Although an equivalence margin was not prespecified for these secondary outcomes a reasonable assumption of equivalence could be made, as the upper and lower bounds of all 95% confidence intervals were within a third of a standard deviation for all analyses (the equivalence limit prespecified for the primary outcome).

Some of the NEPSY-II subscales had large numbers of missing data and the standard deviations were very large with the multiple imputation models. This is because the correlations of the variables included in the multiple imputation model with the outcome variable were low, leading to not much information being recovered using the multiple imputations, while additional noise was added.

Table 6 gives the proportion of children in each group that were reported by a parent to have been diagnosed with a neurodevelopmental disorder and the risk ratio for both as-per-protocol and intention-to-treat analyses. No evidence for any differences was found, with the 95% confidence intervals of all risk ratios crossing 1. However the low event rates limit the inferences that can be drawn regarding equivalence.

The subgroup analyses for the primary outcome are reported in Table 7. These analyses suggest that the differences between groups were similar by age of exposure, and prematurity. Small sample sizes in some of the countries made it inconclusive to interpret country differences in the results. Duration of exposure was not analysed as no children had exposures longer than 120 minutes. The p-value evaluating treatment by country interaction was 0.0496 for the complete case analysis and 0.0643 for the multiple imputation analysis; providing evidence of heterogeneity of the results by country.

In Table 8, the characteristics of children who attended the 5 year follow up are compared to the baseline data of the randomised population and the 2 year outcome data for those who attended

the 2 year follow up. Table 9 demonstrates the unmasking of group allocation for children who attended the 5 year follow up.

Discussion:

In this randomised trial we found strong evidence for equivalence in full scale IQ measured at five years of age between children anaesthetised with awake-regional and general anaesthesia for inguinal herniorraphy in infancy. In a range of other neuropsychological tests evidence of equivalence may also be reasonably assumed as the 95% CI around the differences in means fell within one third of a standard deviation. These results are consistent with the previously reported 2 year outcomes of the GAS trial using the Bayley-III.[16]

The primary outcome was determined at 5 years of age as there is robust evidence for the emergence of the unitary construct of 'general intelligence' and for the individual stability of that construct from middle childhood until adulthood. IQ testing in children around the age 5-6 years has a strong correlation with adult IQ.[23] It has also been shown that IQ aged 5 years is highly predictive of later Maths ability, and that higher IQ in childhood positively predicts a range of benefits in academic, economic and health outcomes across the lifespan.[24] The WPPSI-III is a well-validated, standardised, reliable test for assessing IQ in young children.

~~The IQ, as a measure of intelligence, has significant implications for social-emotional, educational and vocational outcomes throughout the lifespan. The WPPSI-III is an individualised, standardised, reliable and valid test for assessing IQ in young children.~~ The WPPSI-III FSIQ was set as the primary outcome not only due to its strong psychometric properties and predictive potential, but also due to the preclinical data. The widespread cortical damage seen in preclinical models would most likely result in a global decline in function. This would be best identified by a measure of general intellectual function such as the WPPSI-III.

Secondary outcome measures were selected to assess a broad range of cognitive domains that could potentially be impacted based on known vulnerabilities of the developing brain and in response to early animal and human studies. In choosing the tests a number of factors were considered: previous studies found deficits in both hippocampal and non-hippocampal memory; deficits that arise from damage to systems that subserve specific skills are spread through various regions of the brain and are particularly vulnerable to neurological insult (i.e. attention, information processing and executive function); there is a possibility of a cumulative effect of subtle individual or multiple deficits on skill development such as visuo-motor integration,

reading, spelling and arithmetic; and there is previous evidence for social and emotional deficits. Specific individually administered tests and informant report measures were selected from readily available standardised tests in common clinical use with documented reliability and validity statistics for use in this age group.

Several previous cohort studies have sought to identify associations between anaesthesia exposure in early childhood and a range of neurodevelopmental outcomes. The PANDA study was an ambidirectional cohort study that compared neurodevelopmental outcome between children that had previous inguinal herniorraphy and their unexposed siblings using a range of neuropsychological tests performed at 8-15 years of age.[25] This study found no evidence of group differences in IQ scores, or scores on a range of other tests of neurocognitive function and behaviour. Similarly, the MASK cohort study found no evidence for differences between test scores between children that had a single anaesthetic compared to those that had no previous anaesthetics, although children that had multiple anaesthetics did have an increased risk of deficits in processing speed and fine motor outcomes, and parents reported increased problems related to executive function, behaviour and reading.[26] Other cohort studies have found evidence for an association between anaesthesia exposure and cognitive, memory, listening comprehension and language deficits.[27-30]

Several other large population-based data linkage studies have found evidence for an association between anaesthesia in early childhood and a very small decrease in performance in school grades or school readiness tests.[31-34] There is mixed evidence in cohort studies for an association between anaesthesia in early childhood and a subsequent diagnosis of ADHD or other learning disability.[35-42] It is plausible that there may be an increased ~~thc~~ risk of these diagnoses without an increased ~~thc~~ risk of worse outcomes in neurocognitive testing, however other confounding factors are also a possible explanation for these observed associations. The GAS trial found no evidence for an increased risk of behavioural disorders such as ASD or ADHD, however the diagnosis of ADHD and learning disability is typically made in older children, and the low event rate and hence limited power reduced our ability to draw a definitive conclusion.

In all these cohort studies any association found between exposure and poor outcome may be explained by confounding. Children have anaesthesia because they are having surgery or invasive investigations. The condition warranting the procedure may itself be associated with

increased risk of adverse neurodevelopmental outcome. Similarly children with pre-existing but as yet undiagnosed behavioural problems may be at greater risk of needing the procedure. Lastly perioperative factors other than anaesthesia may also increase the risk of poor neurodevelopmental outcome. In most studies, attempts are made to limit the effects of known confounders through patient selection, matching and adjustments in the analysis but the potential influence of confounding can never be eliminated. The GAS trial is the only randomised trial so far that assesses the impact of anaesthesia on neurodevelopment and thus provides the strongest human evidence.

Several previous cohort studies have found more evidence for a detrimental effect after multiple exposures compared to a single exposure. In the GAS trial a substantial number of children had subsequent anaesthetics. The number of children having subsequent anaesthetics was well balanced between arms and thus the occurrence of subsequent anaesthetics is unlikely to influence or bias the results of this trial.

There was weak evidence for an interaction between country and treatment. The reason for this is not immediately apparent and given the marginal level of evidence this finding should be interpreted with caution.

Despite careful selection of patients, an awake-regional technique is not always adequate for herniorraphy. Thus a substantial number of children in the awake-regional group had some exposure to general anaesthetics. These children were excluded in the as-per-protocol analysis. The lack of any substantive difference between the as-per-protocol and intention-to-treat analyses implies that this did not introduce a bias to the trial. In addition, some children were lost to follow up. Multiple imputation was used to reduce the impact of these missing data under the missing at random assumption. However even with multiple imputation the results could be influenced by the selective follow-up of participants. Children who performed poorly at 2 years were more likely to be lost to follow up at 5 years. The reason for this is unclear however this is unlikely to lead to a bias as the 2 year outcome was included in the multiple imputation model. Overall, the loss to follow up was greater than anticipated in the protocol, however the boundaries of the 95% confidence intervals fell within the predefined bounds of equivalence indicating that the precision of the results was adequate in spite of this greater than expected loss to follow up.

Given the nature of the interventions it was impossible to mask the treating surgeons or anaesthetists to group allocation. It was also impractical to completely mask inquisitive parents as adhesives used to secure the airway usually leave signs of skin irritation in the general anaesthesia group, and there would be a puncture mark in the back from the spinal needle in the spinal group. Clinicians making the 5-year assessment were masked successfully in the great majority of cases. It is unlikely that unmasking surgeons, anaesthetists or parents would bias the outcome for the individually administered tests. However, when interpreting parent reported outcomes this potential bias should be considered.

There are considerations to make when assessing the generalisability of the GAS trial. Firstly, the population was predominantly male, which was expected given the surgical pathology selected to create homogeneity within the study sample. Secondly, the infants were exposed over a narrow period of development (early infancy); this period being chosen as the period of high cerebral vulnerability and because this is when both awake regional anaesthesia and general anaesthesia are commonly used for herniorraphy. When determining at which age children might be at greatest risk, it is difficult to translate the animal data to humans.[13,43] In general, younger animals have been found to be at greater risk and thus it would be expected that in humans, infants and the foetus would be most at risk. Some cohort studies have found children exposed at 2-4 years of age to be at greater risk, but this may also be explained by confounding factors, and is less consistent with the preclinical data.[31,32] Thirdly it could be argued that 5 years of age is too early to detect long term neurocognitive outcomes as there are a number of executive functions and social-emotional skills that do not develop until later in life. However these results on individually administered, standardized tests and parent reports indicate that children who undergo anaesthesia in infancy start school life with no neurodevelopmental risk factors. Exploration of executive function and social emotional functions later in development could be an area of future study. Fifthly, in this trial the children received only one general anaesthetic (sevoflurane) in the general anaesthesia group. There are several other general anaesthetics that are used in childhood such as isoflurane, desflurane and propofol. At this stage there are no preclinical data to suggest that any effects seen with sevoflurane would be different to the effects seen with these other agents and thus it is reasonable to assume that the GAS trial results would translate to other general anaesthetic agents. There are also some preclinical data that suggest the effect may be greater if multiple agents are given concurrently. The GAS trial results cannot be

generalized to situations where multiple general anaesthetic agents are given concurrently. Lastly, the length of exposure was on average just under an hour and less than 2 hours for all children. Animal data suggest longer exposures are more likely to cause neurotoxicity, although there is no clear "cut off" for length of exposure that does or does not have an effect. While an hour of anaesthesia was shorter than the exposure used in many of the animal experiments, the equivalence of animal exposure time to that in humans is unknown. Furthermore the median duration of general anaesthesia for children in the 1.5 million procedures in the National Anesthesia Clinical Outcomes Registry (USA) was 57 minutes with infants having a median duration of 79 minutes.[44] Thus the duration of exposure in the GAS trial is longer than nearly half the anaesthetics given to small children.

The number of children potentially affected by national safety warnings about the neurotoxic potential of general anaesthesia such as the FDA warning is significant. During the first 3 years of life approximately 10 percent of children from developed countries will undergo a general anaesthetic for a variety of surgical, diagnostic and medical procedures which translates to millions of children/year.[27,45] Most of these children are healthy and will be exposed to a single short or intermediate length anaesthetic during their childhood.[40] Given the high prevalence of exposure in early childhood, even small effects on brain development due to general anaesthesia could have very large public health consequences. There is also the very real potential that parents and providers will delay necessary procedures in children in an effort to limit exposure at a time of cerebral vulnerability, putting some children at risk for both medical and developmental impairments. The GAS trial, being consistent with data from ~~in addition to~~ several previous cohort studies, provides strong evidence that just under one hour of general anaesthesia in infancy does not cause significant neurocognitive or behavioural deficits.

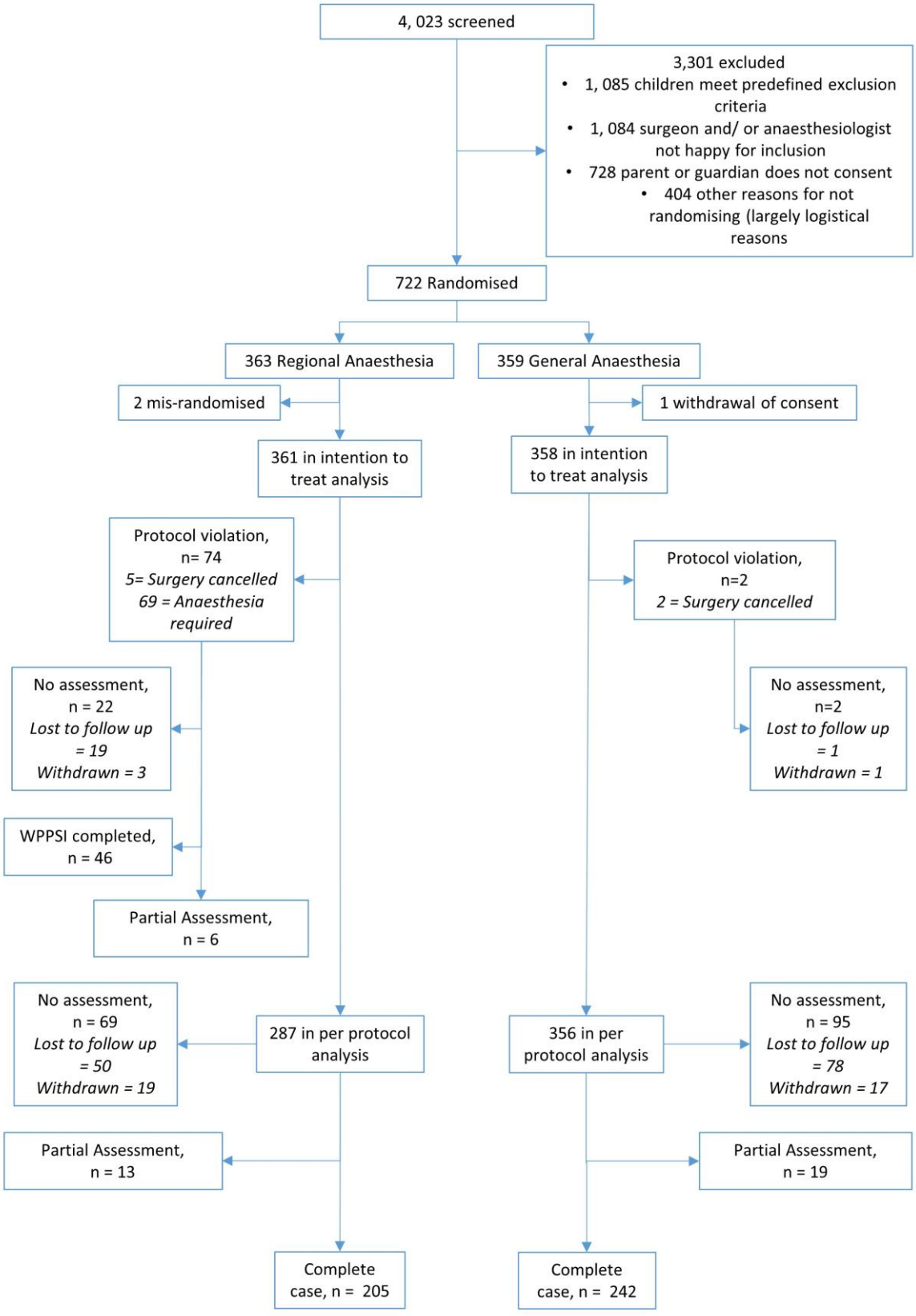
(Word count 5451)

Data sharing statement

The de-identified data set collected for this analysis of the GAS trial will be available six months after publication of this manuscript. The study protocol, analysis plan and consent forms will also be available. The data may be obtained from the Murdoch Children's Research Institute by emailing andrew.davidson@rch.org.au. Prior to releasing any data the following are required: a data access agreement must be signed between relevant parties, the GAS Trial Steering Committee must see and approve the analysis plan describing how the data will be analysed , there must be an agreement around appropriate acknowledgement and any additional costs involved must be covered. Data will only be shared with a recognised research institution which has approved the proposed analysis plan.

588

589 **Figure 1:** Trial Profile



Site	Randomisation RA, N = 361	Randomisation GA, N = 358	Follow-up - complete case* PP - RA, N = 205	Follow-up - complete case* PP - GA, N = 242	Follow-up - complete case* ITT - RA, N = 251	Follow-up - complete case* ITT - GA, N = 242
Australia						
Royal Children's Hospital, Melbourne	57	28	39	44	46	44
Monash Medical Centre, Melbourne	26	25	15	15	16	15
Women's and Children's Hospital, Adelaide	6	5	2	3	4	3
Princess Margaret Hospital for Children, Perth	16	15	7	10	8	10
New Zealand						
Starship Children's Hospital, Auckland	13	12	7	9	8	9
USA						
Children's Hospital, Boston	29	31	18	21	22	21
Children's Memorial Hospital, Chicago	2	3	0	0	0	0
Dartmouth Hitchcock Medical Centre, Lebanon	2	2	2	1	2	1
Vanderbilt Children's Hospital, Nashville	1	2	0	1	0	1
The University of Iowa Hospital, Iowa	8	8	4	4	6	4
Children's Medical Centre, Dallas	7	7	0	7	2	7
Children's Hospital of Philadelphia, Philadelphia	1	1	0	1	0	1
Seattle Children's Hospital, Seattle	11	14	5	13	8	13
The Children's Hospital, Colorado	9	9	4	4	5	4
The University of Vermont/ Fletcher Allen Health Care, Burlington	1	0	1	0	1	0
Canada						
Montreal Children's Hospital, Montreal	21	20	9	15	11	15
Centre de Recherche CHU Sainte-Justine, Montreal	3	5	1	5	3	5
United Kingdom						
Bristol Royal Hospital for Children, Bristol	2	2	1	1	1	1
Royal Hospital for Children, Glasgow	27	25	20	16	21	16
Birmingham Children's Hospital NHS trust, Birmingham	7	6	5	3	6	3
Royal Belfast Hospital for Sick Children, Belfast	2	2	1	0	1	0
Royal Liverpool Children's Hospital, Liverpool	1	1	0	1	0	1
Sheffield Children's Hospital, Sheffield	5	4	0	1	3	1
Italy						
Gaslini Hospital for Children, Genoa	42	39	23	26	30	26
Buzzi Children's Hospital, Milan	25	23	16	15	20	15
Ospedali Riunti, Bergamo	16	20	6	9	7	9
The Netherlands						
Wilhelmina Children's Hospital, University Medical Centre, Utrecht	15	14	14	13	14	13
Universitair Medish Centrum, Groningen	6	5	5	4	6	4

RA = awake-regional anaesthesia. GA = general anaesthesia. PP = per protocol. ITT = intention to treat. *Complete case includes a full WPPSI-III assessed at 5 year follow-up. Results do not include partial assessments.

Table 1: **Enrolment and complete case follow-up by site**

590

	As per protocol		Intention to treat	
	RA group, N = 287	GA group, N = 356	RA group, N = 361	GA group, N = 358
Baseline demographics				
Gender, Male	232 (287, 81%)	304 (356, 85%)	294 (360, 82%)	306 (358, 86%)
Chronological age at surgery (days)	287, 68·9 (31)	356, 71·1 (32)	358, 70·1 (32)	357, 71·0 (32)
Post menstrual age at surgery (days)	287, 317·2 (32)	356, 319·7 (32)	357, 318·3 (33)	357, 319·5 (32)
Weight of child at surgery (kg)	287, 4·2 (1·1)	356, 4·3 (1·1)	359, 4·2 (1·1)	357, 4·3 (1·1)
Pregnancy and birth details				
Mean (SD) Post menstrual age at birth (days)	287, 248·2 (29)	356, 248·6 (27)	360, 248·3 (29)	358, 248·6 (27)
Prematurity (Born < 37 weeks gestation)	160 (287, 56%)	195 (356, 55%)	198 (361, 55%)	196 (358, 55%)
Birth Weight (kg)	287, 2·3 (0·9)	355, 2·3 (0·9)	359, 2·4 (0·9)	357, 2·3 (0·9)
Z score for birth weight	287, -0·7 (1·3)	355, 0·7 (1·3)	359, -0·7 (1·2)	357, -0·7 (1·3)
N, Median (IQR) Apgar score at 1 minute	237, 9 (7-9)	282,8·5 (7-9)	292, 9 (7-9)	284, 9 (7-9)
N, Median (IQR) Apgar score at 5 minutes	237, 9 (9-10)	282, 9 (9-10)	292, 9 (9-10)	284, 9 (9-10)
One of a multiple pregnancy	52 (284, 18%)	61 (356, 17%)	62 (360, 17%)	62 (358, 17%)
Mother received partial course antenatal steroids	16 (287, 6%)	19 (356, 5%)	20 (360, 6%)	19 (358, 5%)
Mother received complete course antenatal steroids	95 (287, 33%)	98 (356, 28%)	114 (360, 32%)	98 (358, 28%)
Mother diagnosed with chorioamnionitis	10 (287, 4%)	12 (356, 3%)	11 (360, 3%)	12 (358, 3%)
Prolonged rupture of the membranes (>24 hours)	28 (287, 10%)	34 (356, 10%)	32 (360, 9%)	34 (358, 10%)
Mother diagnosed with pre-eclampsia	50 (287, 17%)	68 (356, 19%)	60 (360, 17%)	68 (358, 19%)
Sepsis during pregnancy	36 (286, 13%)	50 (356, 14%)	43 (358, 12%)	50 (358, 14%)
Mode of delivery of birth				
Cephalic vaginal	135 (287, 47%)	157 (356, 44%)	169 (360, 47%)	157 (358, 44%)
Breech vaginal	1 (287, <1%)	6 (356, 2%)	3 (360, 1%)	6 (358, 2%)
Compound vaginal	2 (287, 1%)	4 (356, 1%)	3 (360, 1%)	4 (358, 1%)
Caesarean section	149 (287, 52%)	189 (356, 53%)	185 (360, 51%)	191 (358, 53%)
Caesarean section and mother went into labour	42 (287, 15%)	58 (356, 16%)	52 (360, 14%)	59 (358, 16%)
Mother exposed to nitrous oxide during delivery	48 (275, 18%)	62 (344, 18%)	61 (344, 18%)	62 (346, 18%)
IVH	7 (286, 2%)	6 (356, 2%)	8 (359, 2%)	6 (358, 2%)
IVH Grade 1	5 (286, 2%)	6 (356, 2%)	5 (359, 2%)	6 (358, 2%)
IVH Grade 2	2 (286, 1%)	0 (356)	2 (359, 1%)	0 (358)
Retinopathy of prematurity	17 (198, 9%)	16 (256, 6%)	20 (246, 8%)	16 (257, 6%)
Hearing defects detected by perinatal screening	7 (253, 3%)	10 (356, 3%)	8 (316, 3%)	10 (325, 3%)
PDA diagnosed	23 (286, 8%)	21 (355, 6%)	27 (359, 8%)	21 (357, 6%)
PDA never treated	9 (286, 3%)	9 (355, 3%)	11 (359, 3%)	9 (357, 3%)

PDA treated with non-steroidal anti-inflammatory drugs	14 (286, 5%)	10 (355, 3%)	16 (359, 4%)	10 (357, 3%)
Familial Demographics:				
Primary language(s) only spoken*	252 (287, 88%)	305 (356, 86%)	311 (360, 86%)	307 (358, 86%)
Maternal Age at Birth >21	273 (286, 96%)	339 (356, 95%)	339 (358, 95%)	341 (358, 95%)
Family structure two caregivers together, at birth	261 (286, 91%)	324 (356, 91%)	328 (359, 91%)	326 (358, 91%)
Maternal education				
Completed tertiary studies	150 (286, 52%)	171 (354, 48%)	181 (359, 51%)	171 (358, 48%)
Continuing tertiary studies	50 (286, 17%)	67 (354, 19%)	68 (359, 19%)	67 (358, 19%)
Completed year 11 or 12	62 (286, 22%)	83 (354, 23%)	77 (359, 22%)	84 (358, 24%)
Did not complete year 11	25 (286, 9%)	33 (354, 9%)	32 (359, 9%)	34 (358, 10%)
Anaesthesia Details:				
N, Median (IQR)Blood glucose level (mmol/L)	255, 5·4 (4·7-6·1)	314, 5·5 (4·8-6·4)	312, 5·4 (4·7-6·2)	314, 5·5 (4·8-6·4)
Rescue glucose given IV	2 (282, 1%)	4 (356, 1%)	2 (350, 1%)	4 (356, 1%)
Haemoglobin (g/100 ml)	250, 10·3 (2·1)	307, 10·2 (2·0)	305, 10·3 (2·1)	307, 10·2 (2·0)
Need for fluid bolus for hypotension	15 (287, 5%)	59 (356, 17%)	21 (355, 6%)	59 (356, 17%)
Vasoactive drugs given (including atropine)	4 (287, 1%)	17 (356, 5%)	6 (355, 2%)	17 (356, 5%)
N, Median (IQR)Duration of surgery (mins)	286, 26·0 (19·0-35·0)	355, 28·0 (20·0-40·0)	353, 28·0 (20·0-38·0)	355, 28·0 (20·0-40·0)
N, Median (IQR) Duration of sevoflurane exposure (mins)	NA	356, 54·0 (41·0-70·0)	67, 42·0 (31·0-62·5)**	356, 54·0 (41·0-70·0)
Mean end tidal sevoflurane concentration (%)	NA	356, 2·6 (0·7)	67, 2·3 (0·8)**	356, 2·6 (0·7)
Total concentration x hours of exposure	NA	356, 2·6 (1·1)	67, 1·9 (1·0)**	356, 2·6 (1·1)
Any significant apnoea to 12hrs postop***	6 (287, 2%)	15 (356, 4%)	10 (360, 3%)	15 (358, 4%)

Data are n (N, % of non-missing data) or n, mean (SD), unless otherwise stated. APP= As Per Protocol; GA= General Anaesthesia; ITT= Intention to treat; IV= Intra-venously; IVH= Intra ventricular haemorrhage; IQR= Interquartile Range; PDA = Patent ductus arteriosus; RA= Awake Regional Anaesthesia.
* The primary language spoken at home, is the primary language in each country that the Bayley was conducted eg. in Italy it was conducted in Italian
** For those cases that received sevoflurane
*** significant apnoea defined as a pause in breathing for more than 15 seconds or more than 10 seconds if associated with oxygen saturation less than 80% or bradycardia (20% decrease in heart rate)

Table 2: **Baseline descriptive statistics demographic data**

591

	As per protocol		Intention to treat	
	RA group, N = 287	GA group, N = 356	RA group, N = 361	GA group, N = 358
Assessment Details				
Location of 5-year assessment at hospital	198 (216, 91·7%)	228 (257, 88·7%)	246 (268, 91·8%)	228 (257, 88·7%)
Family Demographics at 5 years				

Paid Employment is the main family income		201 (214, 93·9%)	237 (256, 92·6%)	243 (266, 91·4%)	237 (256, 92·6%)
Family Structure, two caregivers living together		194 (214, 90·7%)	223 (257, 86·8%)	230 (266, 86·5%)	223 (257, 86·8%)
Number of children at home					
	1	50 (214, 23·4%)	53 (257, 20·6%)	63 (266, 23·7%)	53 (257, 20·6%)
	2	95 (214, 44·4%)	133 (257, 51·8%)	120 (266, 45·1%)	133 (257, 51·8%)
	3	56 (214, 26·2%)	48 (257, 18·7%)	67 (266, 25·2%)	48 (257, 18·7%)
	> 3	13 (214, 6·1%)	23 (257, 8·9%)	16 (266, 6·0%)	23 (257, 8·9%)
Birth order					
	1	113 (211, 53·6%)	137 (257, 53·3%)	137 (261, 52·5%)	137 (257, 53·3%)
	2	69 (211, 32·7%)	81 (257, 31·5%)	87 (261, 33·3%)	81 (257, 31·5%)
	> 2	29 (211, 13·7%)	39 (257, 15·2%)	37 (261, 14·2%)	39 (257, 15·2%)
Age at follow-up assessment		217, 5·2 (0·2)	258, 5·3 (0·3)	269, 5·2 (0·2)	258, 5·3 (0·3)
Events since original anaesthesia					
Any hospitalisation		101 (199, 50·8%)	129 (250, 51·6%)	131 (249, 52·6%)	129 (250, 51·6%)
Number of days hospitalised					
	0	105 (169, 62·1%)	127 (213, 59·6%)	125 (213, 58·7%)	127 (213, 59·6%)
	1	22 (169, 13·0%)	30 (213, 14·1%)	34 (213, 16·0%)	30 (213, 14·1%)
	2	11 (169, 6·5%)	13 (213, 6·1%)	13 (213, 6·1%)	13 (213, 6·1%)
	>=3	31 (169, 18·3%)	43 (213, 20·2%)	41 (213, 19·2%)	43 (213, 20·2%)
Any anaesthesia		71 (102, 69·6%)	71 (111, 64·0%)	89 (133, 66·9%)	71 (111, 64·0%)
Number of anaesthetics					
	0	104 (156, 66·7%)	132 (181, 72·9%)	131 (197, 66·5%)	134 (183, 73·2%)
	1	28 (156, 17·9%)	27 (181, 14·9%)	37 (197, 18·8%)	27 (183, 14·8%)
	2	11 (156, 7·1%)	11 (181, 6·1%)	14 (197, 7·1%)	11 (183, 6·0%)
	>=3	13 (156, 8·3%)	11 (181, 6·1%)	15 (197, 7·6%)	11 (183, 6·0%)
Any seizures		14 (173, 8·1%)	17 (217, 7·8%)	17 (217, 7·8%)	17 (217, 7·8%)
Events since 2 year assessment					
Child had a head injury that involved loss of consciousness		2 (213, 0·9%)	2 (266, 0·8%)	3 (265, 1·1%)	2 (257, 0·8%)
Child has any chronic illness		38 (213, 17·8%)	43 (258, 16·7%)	48 (265, 18·1%)	43 (258, 16·7%)
Child had any prescribed medication for two months or longer		37 (214, 17·3%)	44 (257, 17·1%)	44 (266, 16·5%)	44 (257, 17·1%)
Child has had an intervention for neurodevelopmental issues		49 (213, 23·0%)	60 (257, 23·3%)	64 (264, 24·2%)	60 (257, 23·3%)
Speech Therapy		36 (217, 16·6%)	48 (259, 18·5%)	50 (269, 18·6%)	48 (259, 18·5%)
Physiotherapy		11 (217, 5·1%)	17 (259, 6·6%)	12 (269, 4·5%)	17 (259, 6·6%)
Occupational Therapy		18 (217, 8·3%)	20 (259, 7·7%)	21 (269, 7·8%)	20 (259, 7·7%)
Psychology		7 (217, 3·2%)	6 (259, 2·3%)	8 (269, 3·0%)	6 (259, 2·3%)
Other interventions		9 (217, 4·1%)	16 (259, 6·2%)	12 (269, 4·5%)	16 (259, 6·2%)
Child attends play group/child care on a regular basis		186 (213, 87·3%)	231 (257, 89·9%)	234 (265, 88·3%)	231 (257, 89·9%)
Physical examination					
Height (cm)		207, 110·8 (5·5)	237, 110·8 (5·5)	254, 110·8 (5·4)	237, 110·8 (5·5)

Weight (kg)	206, 19·3 (3·3)	236, 19·4 (2·8)	253, 19·4 (3·2)	236, 19·4 (2·8)
Head circumference (cm)	194, 51·6 (1·8)	224, 51·2 (2·6)	241, 51·6 (1·8)	224, 51·2 (2·6)
Arm circumference (cm)	191, 17·6 (1·9)	219, 17·4 (1·7)	233, 17·6 (1·9)	219, 17·4 (1·7)

Data are n (N, % of non-missing data) or n, mean (SD). RA = awake-regional anaesthesia. GA = general anaesthesia.

Table 3: **5-year descriptive statistics data**

592

593

594

	APP multiple imputation			APP complete case			ITT multiple imputation			ITT complete case		
	RA group	GA group	Difference in RA-GA*	RA group	GA group	Difference in RA-GA*	RA group	GA group	Difference in RA-GA*	RA group	GA group	Difference in RA-GA*
Global function												
WPPSI III - FSIQ composite score	287, 99·1 (18·4)	356, 99.0 (19·7)	0·2 (-2·6; 3·1)	205, 100·5 (14·3)	242, 100·1 (15·3)	0·6 (-2·1; 3·3)	361, 98·9 (18·0)	358, 98·8 (19·2)	0·2 (-2·5; 2·8)	251, 100·4 (14·1)	242, 100·1 (15·3)	·266 (-2·3; 2·8)
Verbal/language												
WPPSI- III Verbal IQ composite score	287, 100·6 (18·3)	356, 99·7 (20·4)	0·8 (-2·1; 3·8)	206, 101·8 (14·7)	240, 100·9 (15·4)	0·7 (-2·1; 3·4)	361, 99·6 (18·6)	358, 99·6 (19·1)	0·0 (-2·6; 2·7)	251, 101·2 (14·8)	240, 100·9 (15·4)	0·0 (-2·6; 2·5)
NEPSY-II Word Generation scaled score	287, 9·1 (4·7)	356, 9·0 (4·8)	0·1 (-0·6; 0·9)	182, 9·4 (3·4)	199, 9·3 (3·3)	0·1 (-0·6; 0·8)	361, 9·1 (5·5)	358, 9·1 (4·7)	-0·1 (0·6; 0·5)	220, 9·3 (3·5)	199, 9·3 (3·3)	0.1 (-0·6; 0·7)
NEPSY-II Speeded Naming combined scaled score	287, 10·6 (19·6)	356, 7·4 (23·9)	3·3 (-1·1; 7·7)	132, 9·7 (3·0)	142, 9·8 (3·2)	0·0 (-0·7; 0·8)	361, 8·7 (10·3)	358, 9·2 (15·0)	-0·5 (-4·9; 3·9)	162, 9·8 (3·0)	142, 9·8 (3·2)	0·1 (-0·6; 0·8)
Perceptual/visuo-spatial												
WPPSI-III Performance IQ composite score	287, 99·6 (19·3)	356, 100·0 (20·3)	-0·2 (-3·1; 2·8)	206, 100·7 (15·2)	241, 101·2 (15·9)	0·0 (-2·9; 2·8)	361, 100·1 (18·2)	358, 99·8 (19·6)	0·4 (-2·3; 3·1)	252, 101·1 (14·7)	241, 101·2 (15·2)	0·199 (-2·4; 2·8)
NEPSY-II Design Copy scaled score	287, 9·4 (23·8)	356, 6·7 (45·1)	3·1 (-2·7; 8·9)	172, 9·6 (3·4)	207, 9·9 (3·1)	-0.2 (-0·8; 0·5)	361, 13·7 (44·8)	358, 9·6 (26·1)	3·9 (-2·6; 10·4)	212, 9·6 (3·3)	207, 9·9 (3·1)	-0·2 (-0·8; 0·4)
Processing speed												
WPPSI-III Processing Speed Q composite score	287, 95·2 (20·8)	356, 94·7 (21·3)	0·8 (-2·5; 4·0)	196, 95·8 (14·5)	220, 96·3 (15·4)	0·0 (-2·8; 2·9)	361, 95·8 (20·5)	358, 94·6 (21·1)	1·31 (-1·7; 4·3)	241, 96·3 (14·4)	220, 96·3 (15·4)	0·3 (-2·4; 2·9)
Attention/executive function												
NEPSY-II Sentence Repetition scaled score	287, 6·4 (29·7)	356, 8·3 (24·2)	-1·4 (-5·4; 2·7)	175, 9·7 (2·9)	202, 9·7 (2·8)	0·0 (-0·6; 0·6)	361, 13·5 (55·3)	358, 10·8 (23·3)	2·4 (-1·0; 5·8)	214, 9·7 (3.0)	202, 9·7 (2·8)	-0·1 (-0·6; 0·5)
NEPSY-II Auditory Attention combined scaled score	287, 8·7 (4·3)	356, 8·8 (4·6)	-0·1 (-0·8; 0·6)	167, 9·0 (2·7)	183, 9·3 (3·0)	-0·3 (-0·8; 0·3)	361, 8·7 (4·2)	358, 8·8 (5·1)	-0·1 (-0·8; 0·6)	207, 8·9 (3·0)	183, 9·3 (3·0)	-0·3 (-0·8; 0·3)
NEPSY-II Inhibition combined scaled score	287, 7·9 (6·0)	356, 8·4 (5·5)	-0·5 (-1·3; 0·3)	150, 8·3 (3·1)	160, 8·9 (3·0)	-0·6 (-1·3; 0·1)	361, 7·8 (7·2)	358, 8·4 (5·1)	-0·6 (-1·5; 0·4)	179, 8·4 (3·1)	160, 8·9 (3·0)	-0·5 (-1·1; 0·2)
NEPSY-II Statue scaled score	287, 8·6 (33·0)	356, 10·8 (32·1)	-2·6 (-8·9; 3·8)	160, 8·8 (3·5)	182, 8·6 (3·6)	0·2 (-0·5; 1·0)	361, 7·1 (19·3)	358, 8·1 (14·0)	-0·9 (-1·7; -0·2)	192, 8·8 (3·5)	182, 8·6 (3·6)	0·2 (-0·5; 0·9)
CMS Numbers scaled score	287, 8·0 (4·6)	356, 7·8 (4·6)	0·2 (-0·5; 0·9)	194, 8·3 (3·2)	229, 8·1 (3·4)	0·1 (-0·5; 0·7)	361, 7·9 (3·9)	358, 7·7 (4·3)	0·1 (-0·5; 0·8)	236, 8·2 (3·2)	229, 8·1 (3·4)	0·0 (-0·6; 0·6)
Memory & learning												
NEPSY-II Memory for Names combined scaled score	287, 8·1 (4·6)	356, 8·0 (4·6)	0·2 (-0·5; 0·9)	180, 8·1 (3·2)	208, 8·1 (3·2)	0·2 (-0·5; 0·8)	361, 8·2 (4·4)	358, 8·0 (4·6)	0·2 (-0·5; 0·9)	218, 8·2 (3·2)	208, 8·1 (3·2)	0·2 (-0·4; 0·8)
CMS Word Lists I Learning scaled score	287, 8 (4·8)	356, 8·3 (4·9)	-0·4 (-1·1; 0·4)	186, 8·3 (3·4)	224, 8·6 (3·5)	-0·4 (-1·0; 0·3)	361, 8·1 (4·9)	358, 8·3 (5·3)	-0·3 (-1·0; 0·5)	227, 8·3 (3·4)	224, 8·6 (3·5)	-0·3 (-1·0; 0·3)
CMS Word Lists II Delayed scaled score	287, 9·5 (4·0)	356, 9·4 (4·4)	0·1 (-0·5; 0·8)	178, 9·7 (2·8)	209, 9·6 (2·9)	0·0 (-0·5; 0·6)	361, 9·5 (3·9)	358, 9·3 (4·7)	0·1 (-0·5; 0·7)	216, 9·6 (2·9)	209, 9·6 (2·9)	0·0 (-0·6; 0·5)
Social perception												
NEPSY-II Affect Recognition scaled score	287, 10·1 (28·6)	356, 8·9 (18·1)	1·5 (-1·7; 4·6)	174, 10·6 (2·8)	208, 10·4 (3·2)	0·3 (-0·4; 0·9)	361, 11·6 (15·4)	358, 7·4 (74·2)	4·3 (-5·0; 13·5)	215, 10·6 (2·8)	208, 10·4 (3·2)	0·2 (-0·3; 0·8)
NEPSY-II Theory of Mind scaled score	287, 9·3 (4·1)	356, 9·6 (4·6)	-0·3 (-0·9; 0·4)	163, 9·8 (2·9)	178, 9·8 (3·0)	-0·1 (-0·7; 0·5)	361, 9·2 (4·6)	358, 9·6 (4·3)	-0·4 (-1·1; 0·3)	197, 9·7 (3·1)	178, 9·8 (3·1)	-0·2 (-0·8; 0·4)
Sensorimotor												
NEPSY-II Fingertip Tapping Repetitions combined scaled score	287, 9·5 (5·4)	356, 9·4 (5·2)	0.0 (-0·8; 0·8)	180, 9·8 (3·4)	195, 9·7 (3·4)	-0·1 (-0·8; 0·5)	361, 9·6 (4·7)	358, 9·5 (5·3)	0·1 (-0·6; 0·9)	217, 9·8 (3·4)	195, 9·7 (3·4)	0·0 (-0·6; 0·6)
NEPSY-II Fingertip Tapping	287, 7·6	356, 7·1	0·5 (-0·4;	173, 8·1	183, 7·7	0·4 (-0·3;	361, 7·8	358, 7·2	0·6 (-0·4;	204, 8·1	183, 7·7	0·5 (-0·2;

Sequences combined scaled score	(5·3)	(6·6)	1·4)	(3·4)	(3·6)	1·1)	(6·2)	(6·2)	1·6)	(3·4)	(3·6)	1·1)
Academic												
WIAT-II Word Reading composite score	220, 92·1 (20·5)	275, 93·3 (25·9)	-1 (-4·5; 2·5)	147, 92·3 (18·1)	167, 92·8 (21·1)	-1·5 (-4·7; 1·8)	278, 92·1 (23·7)	276, 93·3 (26·6)	-1·2 (-4·6; 2·3)	175, 92·8 (18·8)	167, 92·8 (21·1)	-1·3 (-4·4; 1·8)
WIAT-II Spelling composite score	220, 90·2 (16·3)	275, 91·1 (20·6)	-1·2 (-3·6; 1·2)	141, 90·1 (13·2)	152, 90·8 (16·5)	-1·7 (-4·3; 0·9)	278, 89·9 (17·8)	276, 91·3 (19·2)	-1·6 (-4·2; 1·1)	166, 90·6 (13·7)	152, 90·8 (16·5)	-1·5 (-4·0; 1·0)
WIAT-II Numerical Operations composite score	220, 98·0 (21·3)	275, 96·1 (26·5)	0·8 (-2·8; 4·5)	146, 98·8 (16·2)	161, 96·2 (20·8)	0·3 (-3·1; 3·7)	278, 97·1 (20·8)	276, 96·3 (26·4)	0·5 (-2·9; 3·9)	172, 98·7 (16·6)	161, 96·2 (20·8)	0·2 (-3·0; 3·5)

Data are n, mean (SD). *Difference (95%CI). RA = awake-regional anaesthesia. GA = general anaesthesia. SE = standard error. APP = as per protocol. ITT = intention to treat.

Table 4: **Descriptive statistics WPPSI-III and other individually administered tests for each group**

	APP multiple imputation			**APP complete case**			**ITT multiple imputation**			**ITT complete case**		
	RA group	GA group	Difference in RA-GA*	RA group	GA group	Difference in RA-GA*	RA group	GA group	Difference in RA-GA*	RA group	GA group	Difference in RA-GA*
Executive function												
BRIEF-P (Global Executive composite, T score)	287, 49·2 (16·0)	356, 51·9 (17·6)	-2·7 (-5·2; -0·1)	198, 48·4 (12·5)	232, 51·5 (13·4)	-2·9 (-5·4; -0·4)	361, 49·6 (15·5)	358, 51·9 (17·5)	-2·4 (-4·8; 0·1)	246, 48·9 (12·7)	232, 51·5 (13·4)	-2·4 (-4·7; 0·0)
Adaptive Behaviour												
ABAS-2 (Global Adaptive Behaviour composite score)	287, 94·4 (20·9)	356, 92·6 (23·3)	2·0 (-1·2; 5·2)	168, 95·9 (16·3)	200, 94·1 (16·5)	1·5 (-1·7; 4·8)	361, 94·3 (23·3)	358, 92·5 (23·9)	1·9 (-1·3; 5·1)	205, 95·5 (16·8)	200, 94·1 (16·5)	1·0 (-2·1; 4·2)
Maladaptive Behaviour												
CBCL (Total problems, T score)	287, 45·2 (13·8)	356, 47·1 (16·6)	-2·0 (-4·3; 0·4)	215, 44·6 (11·7)	254, 46·7 (12·5)	-1·9 (-4·1; 0·3)	361, 45·7 (15·0)	358, 47·1 (15·6)	-1·4 (-3·6; 0·8)	265, 45 (12·1)	254, 46·7 (12·5)	-1·5 (-3·6; 0·6)
CBCL (Internalising problems T score)	287, 46·6 (14·4)	356, 48·5 (17·4)	-1·9 (-4·3; 0·6)	215, 46·1 (12·5)	254, 48·0 (12·5)	-1·8 (-4·1; 0·4)	361, 46·8 (15·2)	358, 48·5 (16·0)	-1·6 (-3·9; 0·6)	265, 46·2 (12·5)	254, 48·0 (12·5)	-1·7 (-3·9; 0·4)
CBCL (Externalising problems T score)	287, 44·5 (13·2)	356, 46·1 (15·0)	-1·6 (-3·7; 0·5)	215, 44·0 (10·7)	254, 45·8 (11·9)	-1·7 (-3·7; 0·4)	361, 45·1 (13·9)	358, 46·1 (15·0)	-1·1 (-3·1; 1·0)	265, 44·4 (11·3)	254, 45·8 (11·9)	-1·2 (-3·2; 0·8)

Data are n, mean (SD). *Difference (95%CI). RA = awake-regional anaesthesia. GA = general anaesthesia. SE = standard error. APP = as per protocol. ITT = intention to treat.

Table 5: **Descriptive statistics parent-rated behavioural outcome measures by group**

	As per protocol			Intention to treat		
	RA group, N = 287	GA group, N = 356	RR (95% CI)	RA group, N = 361	GA group, N = 358	RR (95% CI)
Any developmental issues	25 (12·3)	21 (8·8)	1·4 (0·8; 2·4)	33 (12·9)	21 (8·8)	1·5 (0·9; 2·5)
Speech or language issues / interventions	18 (8·4)	17 (6·6)		24 (9)	17 (6·6)	
Psychomotor issues / interventions	8 (3·7)	6 (2·3)		9 (3·4)	6 (2·3)	
Global developmental delay	2 (1)	0 (0)		4 (1·6)	0 (0)	
Behavioural disorders (ADHD, ASD or ODD)	8 (3·8)	15 (6)	0·7 (0·3; 1·7)	13 (4·9)	15 (6)	0·99 (0·5; 2·0)
Diagnosed with Attention Deficit Hyperactivity Disorder (ADHD)	3 (1·4)	4 (1·6)		7 (2·6)	4 (1·6)	
Diagnosed with Autism Spectrum Disorder (ASD)	5 (2·4)	11 (4·4)		7 (2·7)	11 (4·4)	
Hearing abnormality	8 (3·8)	11 (4·4)	0·9 (0·4; 2·2)	12 (4·5)	11 (4·4)	1·1 (0·5; 2·4)
Child has a hearing aid	0 (0)	3 (1·2)		0 (0)	3 (1·2)	
Visual defect of any type in either eye	21 (9·9)	31 (12·2)	0·8 (0·5; 1·3)	28 (10·6)	31 (12·1)	0·8 (0·5; 1·4)
Legally blind (<6/60 in both eyes)	0 (0)	0 (0)		0 (0)	0 (0)	
Cerebral palsy	1 (0·5)	3 (1·2)	0·6 (0·1; 5·5)	1 (0·4)	3 (1·2)	0·4 (0·0; 3·8)

Data are n (% of non-missing data). RR = Risk Ratio. RA = awake-regional anaesthesia. GA = general anaesthesia.

Table 6: **5-year non-psychometric outcome data**

	RA group*	GA group*	Difference in RA-GA	95% CI for difference in RA-GA
Age at surgery (<=70 days)				
APP multiple imputation	111, 98·7 (20·3)	155, 98·2 (19·7)	0·6	-4·1 to 5·3
APP complete case	77, 100·2 (15·1)	107, 99·6 (15·8)	1·0	-3·5 to 5·6
ITT multiple imputation	145, 97·9 (18·6)	155, 98·2 (19·6)	-0·4	-4·8 to 3·9
ITT complete case	97, 99·6 (14·9)	107, 99·6 (15·8)	0·0	-4·2 to 4·2
Age at surgery (>70 days)				
APP multiple imputation	176, 99·7 (17·0)	201, 99·6 (21·0)	0·3	-3·4 to 4·1
APP complete case	128, 100·7 (13·9)	135, 100·5 (14·9)	0·5	-2·9 to 4·0
ITT multiple imputation	213, 100·0 (17·1)	202, 99·6 (18·7)	0·7	-2·7 to 4·1
ITT complete case	152, 100·9 (13·5)	135, 100·5 (14·9)	0·5	-2·7 to 3·8
Australia				
APP multiple imputation	87, 96·0 (16·7)	103, 97·2 (18·4)	-1·2	-6·2 to 3·9
APP complete case	63, 97·7 (13)	72, 98·6 (15·1)	-0·6	-5·4 to 4·3
ITT multiple imputation	105, 96·9 (18·3)	103, 96·8 (18·6)	0·1	-5·2 to 5·3
ITT complete case	74, 98·4 (12·9)	72, 98·6 (15·1)	0·1	-4·5 to 4·6
USA				
APP multiple imputation	49, 99·3 (18·9)	77, 99·6 (19·9)	-0·6	-7·7 to 6·6
APP complete case	34, 101·2 (13·7)	52, 100·2 (16·0)	0·8	-5·6 to 7·2

ITT multiple imputation	71, 98·1 (18·3)	77, 99·5 (18·9)	-1·5	-7·7 to 4·7
ITT complete case	46, 100·2 (13·5)	52, 100·2 (16·0)	-0·5	-6·3 to 5·3
Canada				
APP multiple imputation	16, 93·4 (19·2)	25, 99·1 (19·2)	-5·9	-19·7 to 8·0
APP complete case	10, 97 (12·5)	20, 100·1 (16·9)	-3·8	-17·0 to 9·4
ITT multiple imputation	24, 94·7 (17·1)	25, 99·3 (18·6)	-5·0	-15·4 to 5·5
ITT complete case	14, 95·9 (12·4)	20, 100·1 (16·9)	-4·5	-15·5 to 6·4
New Zealand				
APP multiple imputation	12, 89·5 (19·4)	12, 95·4 (18·5)	-5·9	-23·3 to 11·5
APP complete case	7, 89·6 (13·2)	9, 96·8 (12·9)	-9·9	-24·5 to 4·7
ITT multiple imputation	13, 90·5 (18·2)	12, 96·0 (17·6)	-5·3	20·0 to 9·4
ITT complete case	8, 90·3 (12·4)	9, 96·8 (12·9)	-8·5	-22·5 to 5·4
United Kingdom				
APP multiple imputation	36, 97·8 (19·9)	39, 97·9 (20·1)	1·3	-7·3 to 10·0
APP complete case	27, 98·7 (18·4)	22, 100·1 (15.3)	2·8	-6·5 to 12·0
ITT multiple imputation	44, 96·9 (20·4)	40, 97·6 (20·9)	-0·2	-8·6 to 8·3
ITT complete case	32, 97·8 (18·6)	22, 100·1 (15·3)	0·2	-8·8 to 9·2
Italy				
APP multiple imputation	67, 107·3 (19·0)	81, 101·5 (21·4)	5·6	-1·0 to 12·3
APP complete case	45, 107·8 (12·5)	50, 103·1 (16·2)	4·7	-1·3 to 10·6
ITT multiple imputation	83, 106·5 (17·0)	82, 102·2 (20·9)	4·1	-1·7 to 9·9
ITT complete case	57, 107·2 (11·7)	50, 103·1 (16·2)	4·0	-1·4 to 9·4
The Netherlands				
APP multiple imputation	20, 100·4 (13·4)	19, 99·1 (14·1)	1·3	-7·6 to 10·2
APP complete case	19, 100·3 (12·9)	17, 99·4 (10·3)	1·0	-6·9 to 9·0
ITT multiple imputation	21, 100·9 (13·2)	19, 98·9 (13·1)	2·0	-6·5 to 10·6
ITT complete case	20, 100·7 (12·7)	17, 99·4 (10·3)	1·5	-6·3 to 9·3

*Data are n, M (SD). RA = awake-regional anaesthesia. GA = general anaesthesia. APP = as per protocol. ITT = intention to treat. Note: duration of surgery (< 2 hours vs >= 2 hours) subgroups were not done because all participants had surgery duration < 2 hours

Table 7: **Subgroup analyses for the primary outcome (WPPSI III)**

	Attended 5 year visit			**Did not attend 5 year visit**		
	RA group, N = 271	GA group, N = 259	Total, N = 530	RA group, N = 90	GA group, N = 99	Total, N = 189
Sex of child						
Female	54 (19·9%)	36 (13·9%)	90 (17·0%)	12 (13·5%)	16 (16·2%)	28 (14·9%)
Male	217 (80·1%)	223 (86·1%)	440 (83·0%)	77 (86·5%)	83 (83·8%)	160 (85·1%)
Age (days) at surgery	67·1 (30·2)	71·9 (31·3)	69·5 (30·8)	79·1 (35·0)	68·7 (32·7)	73·6 (34·1)
Birth weight (kg)	2·4 (0·9)	2·3 (0·9)	2·4 (0·9)	2·2 (0·9)	2·3 (0·9)	2·3 (0·9)

Maternal age at birth						
>21	258 (95·2%)	252 (97·3%)	510 (96·2%)	81 (93·1%)	89 (89·9%)	170 (91·4%)
18-21	9 (3·3%)	6 (2·3%)	15 (2·8%)	2 (2·3%)	8 (8·1%)	10 (5·4%)
<18	4 (1·5%)	1 (0·4%)	5 (0·9%)	4 (4·6%)	2 (2·0%)	6 (3·2%)
PMA (days) at birth	249·5 (27·6)	248·4 (27·2)	249·0 (27·4)	244·6 (31·1)	249·0 (27·2)	246·9 (29·1)
Prematurity						
>=37	121 (44·6%)	115 (44·4%)	236 (44·5%)	42 (46·7%)	47 (47·5%)	89 (47·1%)
<37	150 (55·4%)	144 (55·6%)	294 (55·5%)	48 (53·3%)	52 (52·5%)	100 (52·9%)
2-year Bayley-III scores						
Cognitive scaled score	9·9 (2·7)	9·8 (3·0)	9·9 (2·8)	9·2 (3·4)	8·9 (2·6)	9·0 (3·0)
Language composite score	96·2 (14·7)	95·1 (16·0)	95·7 (15·3)	87·8 (17·0)	89·5 (13·3)	88·7 (15·1)
Motor composite score	98·5 (14·2)	97·2 (13·7)	97·9 (13·9)	94·1 (18·4)	96·0 (13·1)	95·1 (15·7)
Social-emotional scaled score	9·8 (3·8)	9·1 (3·6)	9·5 (3·8)	7·9 (3·5)	8·8 (3·7)	8·4 (3·6)
Attended the 2 year visit						
No	55 (20·3%)	45 (17·4%)	100 (18·9%)	44 (48·9%)	43 (43·4%)	87 (46·0%)
Yes	216 (79·7%)	214 (82·6%)	430 (81·1%)	46 (51·1%)	56 (56·6%)	102 (54·0%)

Data are n (%) unless otherwise specified. RA = awake-regional anaesthesia. GA = general anaesthesia. PMA = postmenstrual age

***Table 8*: Characteristics of children that attended the 5 year follow up are compared to the baseline data of the randomised population and the 2 year outcome data for those that attended the 2 year follow up.**

	As per protocol		Intention to treat	
	RA group, N = 287	GA group, N = 356	RA group, N = 361	GA group, N = 358
Psychologist discovered arm of the study the child was randomised to	7 (3.4%)	7 (2.9%)	11 (4.3%)	7 (2.9%)
Paediatrician discovered arm of the study the child was randomised to	13 (8.0%)	13 (6.7%)	16 (7.9%)	13 (6.7%)
Caregiver knew which arm of the study the child was randomised to	105 (51.2%)	118 (47.2%)	131 (51.4%)	118 (47.2%)

Data are n (% of non-missing data). GA= General Anaesthesia; RA= Awake Regional Anaesthesia.

***Table 9:* Details of unmasking at 5 year assessment**

Contributors:

MEMcC was involved in study design, concept and conduct, data coordination, data interpretation, writing the manuscript and revising it critically.

JCdeG was involved in the coordination and supervision of data collection, data analyses and interpretation, revised the manuscript and approved the final manuscript as submitted.

LD contributed to protocol development, data collection, statistical plan, statistical analysis, data interpretation and writing of the manuscript.

ND, DW and GB were involved in study design and conduct, data acquisition and coordination, data interpretation and writing the manuscript.

AG contributed to statistical analyses, statistical analysis plan, data interpretation, and revising the manuscript critically.

RS, RWH, DCB, NM and SS were involved in study design, protocol development, data interpretation and writing the manuscript.

SJS was involved in study conduct, data acquisition and coordination, revising the manuscript and submission of paper.

JM coordinated study conduct in the US including data acquisition and follow-up.

GG was involved in study conduct, data acquisition, data interpretation, editing of paper.

PLH was involved in study conduct, data acquisition, assistance with statistical analysis plan and editing the manuscript critically.

PH was involved in study design, statistical oversight, review of the statistical analysis plan, data interpretation and editing of the manuscript.

GF was involved in study design, data acquisition and contribution to writing the manuscript.

FI, BvS, AL, NW, MM, DP, AA, PS, CB were involved in study conduct, data acquisition and revising the manuscript critically.

AJD was involved in study design and concept, conduct, data coordination, contribution to the statistical analysis plan, data interpretation, writing and coordinating drafts of the manuscript and revising it critically and approving the version to be published.

Declaration of Interests:

We declare no competing interests.

Disclaimer:

The views expressed in this publication are those of the author(s) and not necessarily those of the MRC, NHS, NIHR or the Department of Health UK.

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**: 876-82.

2. Vutskits L, Davidson A. Update on developmental anesthesia neurotoxicity. *Curr Opin Anaesthesiol* 2017; **30**: 337-342.

3. Vutskits L, Xie Z. Lasting impact of general anaesthesia on the brain: mechanisms and relevance. *Nat Rev Neurosci 2016*; **17**: 705-717.

4. Istaphanous GK, Ward CG, Nan X, et al. Characterization and quantification of isoflurane-induced developmental apoptotic cell death in mouse cerebral cortex. *Anesth Analg* 2013; **116**: 845-54.

5. Brambrink AM, Back SA, Riddle A, et al. Isoflurane-induced apoptosis of oligodendrocytes in the neonatal primate brain. *Ann Neurol* 2012; **72**: 525-35.

6. Briner A, De Roo M, Dayer A, et al. Volatile anesthetics rapidly increase dendritic spine density in the rat medial prefrontal cortex during synaptogenesis. *Anesthesiology* 2010; **112**: 546-56.

7. Stratmann G, Sall JW, May LD, et al. Beyond Anesthetic Properties: The Effects of Isoflurane on Brain Cell Death, Neurogenesis, and Long-Term Neurocognitive Function. *Anesth Analg* 2010; **110**: 431-37.

8. Paule MG, Li M, Allen RR, et al. Ketamine anesthesia during the first week of life can cause long-lasting cognitive deficits in rhesus monkeys. *Neurotoxicol Teratol* 2011; **33**: 220-30.

9. Raper J, Alvarado MC, Murphy KL, et al. Multiple Anesthetic Exposure in Infant Monkeys Alters Emotional Reactivity to an Acute Stressor. *Anesthesiology* 2015; **123**: 1084-92.

10. Davidson AJ, Sun LS: Clinical Evidence for Any Effect of Anesthesia on the Developing Brain. *Anesthesiology* 2018; **128**: 840-853.

11. Administration UFaD: FDA Drug Safety Communication: FDA review results in new warnings about using general anesthetics and sedation drugs in young children and pregnant women. 2016.

12. Administration UFaD: FDA Drug Safety Communication: FDA approves label changes for use of general anesthetic and sedation drugs in young children. 2017.

13. Disma N, O'Leary JD, Loepke AW, et al. Anesthesia and the developing brain: A way forward for laboratory and clinical research. *Paediatr Anaesth* 2018; **28**: 758-763.

14. Rappaport BA, Suresh S, Hertz S, et al. Anesthetic neurotoxicity--clinical implications of animal models. *N Engl J Med* 2015; **372**: 796-7.

15. Yahalom B, Athiraman U, Soriano SG, et al. Spinal anesthesia in infant rats: development of a model and assessment of neurologic outcomes. *Anesthesiology* 2011; **114**: 1325-35.

16. Davidson AJ, Disma N, de Graaff JC, et al and the GAS consortium. Neurodevelopmental outcome at 2 years of age after general anaesthesia and awake-regional anaesthesia in infancy (GAS): an international multicentre, randomised controlled trial. *Lancet* 2016; **387**: 239-50.

17. Davidson AJ, Morton NS, Arnup SJ, et al and the GAS consortium: Apnea after Awake Regional and General Anesthesia in Infants: The General Anesthesia Compared to Spinal Anesthesia Study--Comparing Apnea and Neurodevelopmental Outcomes, a Randomized Controlled Trial. *Anesthesiology* 2015; **123**: 38-54.

18. McCann ME, Withington DE, Arnup SJ, et al the GAS Consortium. Differences in Blood Pressure in Infants After General Anesthesia Compared to Awake Regional Anesthesia (GAS Study-A Prospective Randomized Trial). *Anesth Analg* 2017; **125**: 837-845.

19. Frawley G, Bell G, Disma N, et al and the GAS consortium. Predictors of Failure of Awake Regional Anesthesia for Neonatal Hernia Repair: Data from the General Anesthesia Compared to Spinal Anesthesia Study--Comparing Apnea and Neurodevelopmental Outcomes. *Anesthesiology* 2015; **123**: 55-65.

20. Disma N, Withington D, McCann ME, et al and the GAS Consortium: Surgical practice and outcome in 711 neonates and infants undergoing hernia repair in a large multicenter RCT: Secondary results from the GAS Study. *J Pediatr Surg* 2018; **53**: 1643-1650.

21. Davidson A, McCann ME, Morton N et al. www.thelancet.com/protocol-reviews/09prt-9078:

22. Piaggio G, Elbourne DR, Altman DG, Pocock SJ, Evans SJW, for the CONSORT Group. Reporting of Noninferiority and Equivalence Randomized Trials: An Extension of the CONSORT Statement. *JAMA* 2006; **295**: 1152-1160.

23. Hindley CB, Owen CF. The extent of individual changes in IQ for ages between 6 months and 17 years, in a British longitudinal sample. *Journal of Child Psychology and Psychiatry* 1978; **19:** 329–350.

24. Batty GD, Der G, Macintyre S, Deary IJ. Does IQ explain socioeconomic inequalities in health? Evidence from a population based cohort study in the west of Scotland. *BMJ* 2006; **332** :580.

25. 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.

26. 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.

27. Ing C, DiMaggio C, Whitehouse A, et al. Long-term differences in language and cognitive function after childhood exposure to anesthesia. *Pediatrics* 2012; **130**: e476-85.

28. 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.

29. Backeljauw B, Holland SK, Altaye M, et al. Cognition and Brain Structure Following Early Childhood Surgery With Anesthesia. *Pediatrics* 2015; **136**: e1-12.

30. de Heer IJ, Tiemeier H, Hoeks SE, et al. Intelligence quotient scores at the age of 6 years in children anaesthetised before the age of 5 years. *Anaesthesia* 2017; **72**: 57-62.

31. Graham MR, Brownell M, Chateau DG, et al. Neurodevelopmental Assessment in Kindergarten in Children Exposed to General Anesthesia before the Age of 4 Years: A Retrospective Matched Cohort Study. *Anesthesiology* 2016; **125**: 667-677.

32. 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.

33. Glatz P, Sandin RH, Pedersen NL, et al. Association of Anesthesia and Surgery During Childhood With Long-term Academic Performance. *JAMA Pediatr* 2017; **171**: e163470.

34. Clausen NG, Pedersen DA, Pedersen JK, et al. Oral Clefts and Academic Performance in Adolescence: The Impact of Anesthesia-Related Neurotoxicity, Timing of Surgery, and Type of Oral Clefts. *Cleft Palate Craniofac J* 2017; **54**: 371-380.

35. DiMaggio C, Sun LS, Kakavouli A, Byrne MW, Li G: A retrospective cohort study of the association of anesthesia and hernia repair surgery with behavioral and developmental disorders in young children. J Neurosurg Anesthesiol 2009; 21: 286-91.

36. DiMaggio C, Sun LS, Li G. Early childhood exposure to anesthesia and risk of developmental and behavioral disorders in a sibling birth cohort. *Anesth Analg* 2011; **113**: 1143-51.

37. Flick RP, Katusic SK, Colligan RC, et al. Cognitive and behavioral outcomes after early exposure to anesthesia and surgery. *Pediatrics* 2011; **128**: e1053-61.

38. Sprung J, Flick RP, Katusic SK, et al. Attention-deficit/hyperactivity disorder after early exposure to procedures requiring general anesthesia. *Mayo Clin Proc* 2012; **87**: 120-9.

39. Bong CL, Allen JC, Kim JT. The effects of exposure to general anesthesia in infancy on academic performance at age 12. *Anesth Analg* 2013; **117**: 1419-28.

40. 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**: 227-240.

41. 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. *Paediatr Anaesth* 2014; **24**: 741-8.

42. 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. *Eur J Anaesthesiol* 2015; **32**: 303-10.

43. Montana MC, Evers AS: Anesthetic Neurotoxicity: New Findings and Future Directions. *J Pediatr* 2017; **181**: 279-285.

44. Bartels DD, McCann ME, Davidson AJ, et al. Estimating pediatric general anesthesia exposure: Quantifying duration and risk. *Paediatr Anaesth* 2018; **28**: 520-527.

45. 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.

GAS Study Consortium

AUSTRALIA

Andrew J. Davidson and Geoff Frawley (Department of Anaesthesia and Pain Management, Murdoch Children's Research Institute and The Royal Children's Hospital and University of Melbourne, Melbourne, Australia); Pollyanna Hardy (Birmingham Clinical Trials Unit, University of Birmingham, UK); Sarah J. Arnup, Anneke Grobler and Katherine Lee (Clinical Epidemiology and Biostatistics Unit, Murdoch Children's Research Institute, Melbourne, Australia); Rodney W. Hunt (Department of Neonatal Medicine, The Royal Children's Hospital and Murdoch Children's Research Institute and University of Melbourne, Melbourne, Australia); Robyn Stargatt (School of Psychological Science, La Trobe University and Child Neuropsychology, Murdoch Children's Research Institute, Melbourne, Australia); Suzette J. Sheppard, Gillian D. Ormond, Penelope L. Hartmann, Michael J. Takagi, Kaitlyn Taylor, Stephanie Malarbi and Melissa Doyle (Department of Anaesthesia and Pain Management, Murdoch Children's Research Institute, Melbourne, Australia); Philip Ragg (Department of Anaesthesia and Pain Management, The Royal Children's Hospital, Melbourne, Australia); David Costi (Paediatric Anaesthesia, Women's and Children's Hospital, Adelaide, Australia); Britta S. von Ungern-Sternberg (Department of Anaesthesia and Pain Management, Princess Margaret Hospital for Children and The University of Western Australia, Perth, Australia)

NZ

Niall C. Wilton, and Graham Knottenbelt (Department of Paediatric Anaesthesia and Operating Rooms, Starship Children's Hospital, Auckland, New Zealand)

CANADA

Davinia Withington (Department of Anesthesia, Montreal Children's Hospital and McGill University, Montreal, Canada); Koto Furue, Hélène Gagnon (Département d'Anesthésie, Centre Hospitalier Universitaire Sainte-Justine, Montreal, Canada)

ITALY

Nicola Disma, Leila Mameli and Gaia Giribaldi (Department of Anesthesia, Istituto Giannina Gaslini, Genoa, Italy); Alessio Pini Prato (Department of Pediatric Surgery, The Children Hospital, AON SS Antonio e Biagio e Cesare Arrigo, Alessandria, Italy); Girolamo Mattioli (DINOGMI University of Genoa, Genoa, Italy); Andrea Wolfler and Francesca Izzo (Pediatric Anesthesia and Intensive Care Unit, Department of Pediatrics, Ospedale dei Bambini V Buzzi, ASST Fatebenefratelli Sacco, Milan, Italy); Stefania Maria Bova, Arianna Krachmalnicoff, Child Neurology Unit, Ospedale dei Bambini V Buzzi, ASST Fatebenefratelli Sacco, University of Milan, Milan, Italy; Claudia Guuva (Unit of Child Neurology and Psychiatry, ASST Papa Giovanni XXXIII Bergamo, Italy)

NETHERLANDS

Jurgen C. de Graaff (Department of Anesthesiology, Erasmus Medical Centre, Rotterdam, The Netherlands and Department of Anesthesiology, University Medical Centre Utrecht, Utrecht University, The Netherlands); Desiree B.M. van der Werff, Jose T.D.G van Gool, Kim van Loon and Cor J. Kalkman (Department of Anesthesiology, University Medical Centre Utrecht, Utrecht University, The Netherlands); Anneloes L. van Baar (Utrecht Centre for Child and Adolescent Studies. Utrecht University, Utrecht, The Netherlands); Anthony R. Absalom, Frouckje M. Hoekstra, Martin Volkers and Martine Oostra (Department of Anesthesiology, University Medical Center Groningen, Groningen University, Groningen, The Netherlands)

UK

Graham Bell (Department of Anaesthesia, Royal Hospital for Children, Glasgow, Scotland, UK); Liam Dorris (Paediatric Neurosciences, Royal Hospital for Children and Institute of Health and Wellbeing, University of Glasgow, Scotland, UK); Neil S. Morton (University of Glasgow and Department of Anaesthesia, Royal Hospital for Children, Glasgow, Scotland, UK); Jaycee Pownall and Jack Waldman (Institute of Health and Wellbeing, University of Glasgow, Scotland, UK) Ruth Hind, Joseph D Symonds (Paediatric Neurosciences, Royal Hospital for Children, Glasgow, Scotland, UK); Oliver Bagshaw (Anaesthetic Department, Birmingham Children's Hospital, Birmingham, UK)

US

Mary Ellen McCann, Charles Berde, Sulpicio Soriano, Navil Sethna, Pete Kovatsis, and Joseph Cravero (Department of Anesthesiology, Critical Care and Pain Medicine, Children's Hospital Boston, Boston, USA); David Bellinger and Jacki Marmor (Department of Neurology, Children's Hospital Boston, Boston, USA); Anne Lynn, Iskra Ivanova, Agnes Hunyady, and Shilpa Verma (University of Washington, Seattle Children's Hospital, Department of Anesthesia and Pain Medicine, Seattle, USA); David M. Polaner (Children's Hospital Colorado and University of Colorado, Department of Anesthesiology, Colorado, USA); Joss Thomas, Martin Mueller, and Denisa Haret (The University of Iowa Hospital, Department of Anesthesia, Iowa, USA); Peter Szmuk, Jeffrey Steiner, Brian Kravitz and Alan Farrow-Gillespie (Children's Medical Centre Dallas, Department of Anesthesiology, University of Texas Southwestern Medical Center, Dallas and Children's Medical Center at Dallas and Outcome Research Consortium, Texas, USA); Santhanam Suresh (Department of Pediatric Anesthesiology, Ann & Robert H Lurie Children's Hospital of Chicago, Illinois, USA); Stephen Hays (Pediatric Anesthesia, Monroe Carell Jr. Children's Hospital at Vanderbilt, Tennessee, USA); Andreas Taenzer (Dartmouth-Hitchcock Medical Center, Department of Anesthesiology, New Hampshire, USA); Lynne Maxwell (Department of Anesthesiology and Critical Care Children's Hospital of Philadelphia, Philadelphia, USA); Robert K. Williams (Anesthesia and Pediatrics, College of Medicine, University of Vermont, Vermont Children's Hospital, Vermont, USA)