



Published in final edited form as:

Contemp Clin Trials. 2015 March ; 41: 45–54. doi:10.1016/j.cct.2014.12.020.

Neurodevelopment of Children Exposed to Anesthesia: Design of the Mayo Anesthesia Safety in Kids (MASK) Study

Stephen J. Gleich, M.D.¹, Randall Flick, M.D., M.P.H.¹, Danqing Hu, M.D.², Michael J. Zaccariello, Ph.D., L.P.³, Robert C. Colligan, Ph.D., L.P.³, Slavica K. Katusic, M.D.⁴, Darrell R. Schroeder, M.S.⁴, Andrew Hanson, B.S.⁴, Shonie Buenvenida, R.N.¹, Robert T. Wilder, M.D., Ph.D.¹, Juraj Sprung, M.D., Ph.D.¹, Robert G. Voigt, M.D.⁵, Merle G. Paule, Ph.D.⁶, John J. Chelonis, Ph.D.⁶, and David O. Warner, M.D.¹

¹Department of Anesthesiology, Mayo Clinic, Rochester, Minnesota

²Mayo Graduate School, Mayo Clinic, Rochester, Minnesota

³Department of Psychology, Mayo Clinic, Rochester, Minnesota

⁴Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota

⁵Department of Pediatric Medicine, Baylor College of Medicine, Houston, Texas

⁶National Center for Toxicological Research, Food and Drug Administration, Jefferson, Arkansas

Abstract

There is increasing evidence that exposure of developing brains in animals, including nonhuman primates, to commonly-utilized anesthetic agents may cause adverse effects on cognition and behavior. In this paper, we summarize our methodology for a population-based, propensity-matched study to evaluate possible anesthesia-related sequelae in preschool children when evaluated in elementary or high school. A cohort of all children born in Olmsted County, Minnesota between the years 1994-2007 who are currently local residents has been identified. Existing medical records are being used to identify all episodes of exposure to general anesthesia prior to the age of 3 years (i.e., prior to their 3rd birthday). Children with multiple, single, and no anesthesia exposure are sampled for testing between the ages of 8-12 years or 15-19 years during the period 2012-2016. To match children in different exposure groups as closely as possible, sampling is guided by propensity-matching for the likelihood of receiving anesthesia. Selected children are invited to participate in a single 4-hour session of neuropsychological testing, including the National Center for Toxicological Research-Operant Test Battery, which has been used to study anesthetic neurotoxicity in nonhuman primates. The results of this testing will be compared among children with different anesthetic exposure histories. The expected products of this research will be a detailed phenotype of possible anesthetic-associated neurotoxicity in

© 2014 Elsevier Inc. All rights reserved.

Corresponding Author: David O. Warner, M.D., Mayo Clinic, Department of Anesthesiology, 200 1st St SW, Rochester, MN 55905, Ph: 507-255-4288, warner.david@mayo.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

humans, utilizing a robust patient database and neuropsychological testing battery, and the first comparison of effects of anesthetic exposure in children and nonhuman primates performing nearly identical behavioral tasks.

1. Introduction

There is mounting evidence that exposure of the developing brains of rodents and nonhuman primates to anesthetic and sedative medications may cause neurodegenerative changes¹⁻⁵ with subsequent adverse effects on learning and behavior.⁶⁻¹⁰ For example, exposure of neonatal rhesus monkeys to ketamine causes both neuroapoptosis and long-term decrements in performance on specific tests of cognition and behavior, including the National Center for Toxicological Research-Operant Test Battery (NCTR-OTB).¹¹ Implicated drugs, among animal studies, include N-methyl-D-aspartate (NMDA) glutamate receptor antagonists (e.g. ketamine and nitrous oxide) and agents with gamma-aminobutyric acid-mimetic (GABA_A) properties (e.g. midazolam, isoflurane, halothane, and propofol) at doses comparable to those used in human clinical practice. As more than one million young children are exposed to drugs used to induce and maintain general anesthesia each year in the United States,¹² the clinical implications of potential anesthesia-related neurotoxicity in humans are profound.

Several observational studies have examined the relationship between exposure to anesthesia/surgery and neurodevelopmental outcomes in humans. In general, many, but not all studies have found associations between such exposure and adverse outcomes.¹³⁻¹⁸ These include learning disabilities, attention-deficit hyperactivity disorder (ADHD), impaired performance on standardized achievement tests, and measures of memory. For example, in a birth cohort of children born in Olmsted County from 1976-1982, multiple, but not single, exposures to anesthesia prior to age 3 were associated with the subsequent development of learning disabilities and ADHD.^{14,15} Indeed, multiple anesthetic exposures were associated with approximately a doubling in the incidence of learning disabilities. However, existing epidemiologic measures readily available may not be appropriate instruments to describe anesthesia-associated sequelae and cannot be directly compared to behavioral measures such as those obtained using the NCTR-OTB in nonhuman primates. Thus, among humans, a detailed phenotype of any anesthesia-associated sequelae remains to be defined.

This manuscript describes the methodology for the ongoing Mayo Anesthesia Safety in Kids (MASK) study. This project will sample 1,000 children from a population-based birth cohort for detailed neurodevelopmental assessments of behavior and learning, including those obtained from the NCTR-OTB. Children exposed to anesthesia prior to age 3 years and a reference sample of propensity-matched children not exposed to anesthesia are being evaluated. The objective is to determine whether exposure of children to anesthesia prior to age 3 is associated with neurodevelopmental abnormalities. Our results will provide a detailed phenotype of anesthetic-associated outcomes, if any, and will be the first direct comparison of the effects of anesthetic exposure on performance of the NCTR-OTB in animals and children/adolescents.

2. Materials and methods

2.1 Hypotheses

This study tests the following hypotheses:

2.1.1 Primary hypothesis—Exposure of children to multiple anesthetics prior to age 3 years is associated with impaired performance on specific domains of neurocognition.

2.1.2 Secondary hypothesis—The pattern of impairments on performance of the behavioral tasks that make up the NCTR-OTB will be similar in rhesus monkeys (historical data) and multiply-exposed children.

2.2 Study design overview

A cohort of all children born in Olmsted County, Minnesota between the years 1994-2007 who are currently local residents has been identified. Existing medical records are being used to identify all episodes of exposure to general anesthesia prior to the age of 3 years (i.e., prior to their 3rd birthday). Preschool children with multiple, single, and no anesthesia exposures are being selected for evaluation between the ages of 8-12 years or 15-19 years during 2012-2016. Sampling is guided by propensity-matching for the likelihood of receiving anesthesia. Selected children are being invited to participate in a single 4-hour session of neuropsychological testing, which includes performance of the NCTR-OTB. These results will be compared among children with different anesthetic exposure histories.

2.3. Setting

This study includes children born in Olmsted County, Minnesota. According to U.S. census data in 2000,¹⁹ the Olmsted County population is comparable to the overall U.S. population with respect to median age (35.0 years vs. 35.3 years), median household income (\$51,316 vs. \$42,000), and male gender (49.1% vs. 49.1%). The percent of white population is higher than in the U.S. (90.3% vs. 75.1%) when all residents are considered. However, the community is rapidly becoming more diverse as indicated by the racial and ethnic characteristics of children enrolled in public schools in Rochester. In 2007, of the 16,330 children enrolled, 27% were non-white, with 11% black, 7% Hispanic, and 9% Asian.

2.4 Creation of population-based birth cohort

Using 1) the resources of the Rochester Epidemiology Project (REP),²⁰ a medical records linkage system that provides access to the complete medical records of all Olmsted County residents, and 2) birth certificate information from the Minnesota Department of Health, Division of Vital Statistics, all children born from January 1, 1994 to December 31, 2007 to mothers residing in Olmsted County at the time of the child's birth were identified. This date range was chosen as 1) coinciding with the introduction of the modern anesthetic sevoflurane into clinical practice, which rapidly replaced halothane, and 2) ending with 2007 as a date that would provide children who could be tested at age 8 or greater during the study period. Birth certificate information included the child's name, sex, and date of birth, the mother's name, the father's name, and the mother's address at the time of the child's birth. The population targeted for recruitment consists of those children in this birth cohort who 1)

resided within Olmsted County up to the age of 3 (to ensure complete ascertainment of anesthesia records) and 2) resided within 25 miles of Rochester, Minnesota (in the vicinity of Olmsted County) during the recruitment period (to enhance the feasibility to return for testing). The potential influence of migration bias (moved or deceased before age 8) will be assessed during study analysis using birth certificates and/or medical records (e.g., comorbidities). As described above and in prior publications,²¹ we have not found migration bias to be a significant factor in studies to date using similar Olmsted County birth cohorts.

2.5. Target enrollment

One group of children is being evaluated between the ages of 8-12 years (during the elementary school years), and another between the ages of 15-19 years (the high school years). These ages were chosen to provide a sufficient number of children for evaluation, as representing two developmental stages (pre-adolescence and adolescence) over which the results of testing would be comparable. Children with severe mental retardation are excluded, as these children would most likely be unable to complete the detailed neuropsychological assessment and perform the NCTR-OTB used in this study.

Factors determining the target number of children to be recruited in each age group (Figure 1) included: 1) preliminary estimates of the number of children who would be exposed to anesthesia based on our prior work; 2) considerations of statistical power (see section 2.12.1); 3) the feasibility of recruiting busy school-aged children (and their parents) to a half-day testing session; 4) capacity of our Psychological Assessment Lab; and 5) logistical challenges of scheduling appointments that meet the needs of children/adolescents and their families. Based these factors, we plan on testing on average one child per working day over the course of the testing period from December 2012 to December 2016 (4 years), or 250 children per year. Additionally, weekend appointments are offered allowing two more children to be assessed within a given week. Sampling and recruitment of more children with single anesthetic exposures is important for two reasons: 1) greater availability of children with single exposures; and 2) effects of single exposures may be less than those of multiple exposures, thus we want to increase our statistical power to detect subtle differences. Based on conservative estimates performed at the time of initial study planning of the number of children who would be exposed to anesthesia, our target enrollment requires us to recruit approximately half of all children born in Olmsted County who were exposed to multiple anesthetics for testing and a quarter of children exposed to single anesthetics. As the majority (>90%) of children are not exposed to anesthesia prior to age 3 years, we do not anticipate any challenge with recruiting children who have not been exposed to anesthesia.

2.6. Selection of children for recruitment

Because a high proportion of available multiply-exposed children will need to be recruited, all eligible multiply-exposed children are being contacted (see section 2.7). Because of the potential for confounding (see section 3.2), single- and non-exposed children are selected for recruitment using a frequency-matched approach with strata defined based on their propensity for receiving general anesthesia, a method similar to that utilized in our prior

work.¹⁴ Conceptually, this method attempts to select singly- and non-exposed children who are best matched on a variety of characteristics to those children who are multiply-exposed.

2.6.1. Propensity matching—The health status of each cohort member eligible for recruitment is quantified using the Johns Hopkins ACG Case-Mix System²² which utilizes International Classification of Diseases; Ninth Revision diagnosis codes (ICD 9). Codes for each child entered up to and during the third year of life are assigned to one of 32 unique morbidity clusters designated as aggregated diagnostic groups (ADG) based on clinical criteria including duration of care, severity, diagnostic certainty, type of etiology, and expected need for subspecialty care. Multinomial logistic regression is used to calculate the propensity for receiving general anesthesia.

For this analysis, the dependent variable is anesthesia exposure (none, single, multiple). Separate propensity models are fit for males and females with 32 binary indicator variables included representing the ACG morbidity clusters as well as variables available from the birth certificate or other REP resources including gestational age at birth, birth weight, APGAR score, mother's and father's age and level of education, etc. From each of these multinomial logistic regression models we obtain a propensity score for receiving a single exposure to general anesthesia and a propensity score for receiving multiple exposures to general anesthesia. For both propensity scores (single exposure and multiple exposure), a categorical variable is created based on the quintiles of the observed distribution. Using all possible combinations of these 2 categorical variables results in 25 propensity matched strata. For each testing age range, the goal is to include a total of 500 subjects (100 subjects with multiple exposures to general anesthesia, 150 subjects with a single exposure to general anesthesia and 250 subjects with no exposure to general anesthesia). Within each of the propensity matched strata, subjects from each of the exposure groups (single- and non-exposed) are randomly selected for potential inclusion in the study.

2.6.2. Sampling—In order to obtain the appropriate frequency matching within a given strata (Figure 1), for every 2 subjects enrolled with multiple exposures to general anesthesia, 3 subjects are enrolled with one exposure and 5 subjects with no exposure. Since it is not known how many of those randomly selected will agree to participate, the recruitment of study subjects is performed in batches (see section 2.7) with the resulting sample-sizes monitored and appropriate sampling adjustments made to ensure that appropriate frequency matching of exposed and unexposed individuals is achieved within each strata. We have chosen not to attempt to use siblings for matching purposes, as 1) health status or other potential determinants of neurodevelopment may differ among siblings; and 2) we want to find a match for as many exposed children as possible, and siblings may not be available for all exposed children.

2.7. Recruitment

Batches of invitation letters are being sent at intervals to the families of potential study participants and include a brochure explaining study purposes and procedures. Letters invite the recipient to respond via mail or phone call if interested or wanting more information. Telephone calls are made by study personnel to non-responders and those indicating interest.

In addition to letters of invitation, those children multiply-exposed to anesthetics are specifically targeted for recruitment in a series of television and newspaper advertisements inviting them to participate and to call study personnel for more information.

2.8. Consent procedures and remuneration

For children to be tested who are 8-17 years old, IRB-approved consent is obtained from a parent (or guardian), and IRB-approved assent is obtained from the child. For participants to be tested who are adults (age 18-19), IRB-approved consent is obtained from these individuals and a separate consent is obtained from their parent or guardian to complete parental questionnaires. Consent is obtained on the day of testing and each child or adolescent participant receives \$100 remuneration.

2.9 Study measurements

Three categories of assessments are being obtained:

1. Subject information, including anesthesia and birth certificate information
2. A Neuropsychological Assessment Battery utilizing established measures of cognitive and behavioral functions
3. Performance in the NCTR-OTB

The latter two assessments are obtained during a single approximately 4-hour testing session conducted by a trained psychometrist at the Psychological Assessment Laboratory at Mayo Clinic in Rochester, Minnesota and supervised by a board certified clinical neuropsychologist.

2.9.1. Anesthesia and birth certificate information

Anesthesia/surgical information: For those children exposed to general anesthesia prior to their 3rd birthday, the following information is abstracted by study personnel from the medical records for each anesthetic exposure: American Society of Anesthesiologists physical status classification, type of surgery or procedure and urgency, total duration of anesthesia, number of anesthetic exposures, dates on which exposure occurred, and anesthetic agents used.

Birth certificate information: For all children, the following information is available from the birth certificates: date of birth, sex, gestational age at birth, birth weight, APGAR scores at 1 and 5 min, complications of pregnancy, complications of labor and delivery, the number of births (multiple or single), need for induced labor, and mother's and father's age and level of education (<12 years, 12 years, >12 years).

2.9.2. Neuropsychological Assessment Battery—The goal is to assess cognitive, attention/memory, executive-functioning, fine motor development, and behavioral characteristics, using neuropsychological assessment measures that are well-established, state-of-the-art, psychometrically sound, and feasible within an approximately 3-hour testing period (Table 1).²³⁻²⁵ Observational checklists of social/behavioral/academic functioning from suitable external sources, i.e., parents/guardians, will also be utilized.

2.9.2.1. Cognitive/intellectual abilities: The Wechsler Abbreviated Scale of Intelligence (WASI)²⁶ was developed to be a brief, reliable and valid index of verbal, non-verbal, and general intellectual abilities, strongly related to, and modeled after, the child and adult versions of the Wechsler Intelligence Scales and intended to be applicable in research settings. A two-subtest version of the WASI is administered: 1) *Vocabulary*: the 42 items are presented in pictures, words and verbal prompts, with the participant required to define each word. Vocabulary (defining words) is a good measure of general intelligence, expressive vocabulary and fund-of-knowledge (crystallized intelligence); 2) *Matrix Reasoning*: the 35 incomplete matrices (i.e., patterns superimposed on a grid) have one piece missing, with the correct response (i.e., the missing piece) to be selected from 5 alternatives (a multiple-choice format). This non-verbal task is also a measure of general intelligence, and non-verbal reasoning (fluid intelligence). An overall Full-Scale IQ standard score is calculated. Additionally, a Verbal IQ (T-score) and Performance IQ (T-score) quotient is derived.

2.9.2.2. Memory: The following tasks of the Wide Range Assessment of Memory and Learning-2nd Edition (WRAML-2)²⁷ are administered. (1) *Story-Memory*: an age-appropriate story of 8-10 sentences is read to the participant and they are asked to remember as many details as they are able. Delayed and recognition-recall scores are also obtained. Immediate learning, delayed recall and recognition are variables of interest. (2) *Verbal Learning*: a list of 13-16 words is read aloud to the participant. After each of four training trials, the number of words recalled is recorded. Delayed and recognition-recall scores are also obtained. Immediate learning, delayed recall and recognition scaled scores will be calculated. (3) *Design Memory*: one at a time, five cards carrying simple line drawings, of increasing complexity, are shown to the participant, for 5 seconds. For each card, after a 10-second delay, the participant is asked to reproduce/draw the design from memory. Later, portions of the designs and foils are presented visually, and the participant is asked to recall if the specific stimuli was presented during the learning phase. Immediate learning and delayed recognition scores are the variables of interest. These 3 tasks were chosen from the total WRAML-2 subtests as representative of significant core domains of this assessment, including immediate recall for just-presented material, and retention/recall of verbal or visual information.

2.9.2.3. Attention: The Conners Continuous Performance Test (2nd Edition)²⁸ is a measure of the participant's ability to sustain attention for an extended period of time and respond to an identified stimulus. This instrument also measures a participant's ability to inhibit incorrect responses. Scores of interest are: number of omissions, number of commissions and reaction time. In addition, three additional subtests from the WRAML-2²⁷ will also assess attention: (1) *Finger Windows*: this is a nonverbal working memory task relying on a stimulus card of asymmetrical holes. The examiner first demonstrates, with his/her finger increasingly longer strings of sequences. The participant then mimics the examiner; (2) *Number-Letter*: the participant is asked to repeat verbatim, an increasingly longer string of numbers and letters in correct order; (3) *Sentence Memory*: this task will assess the participant's ability to repeat a series of sentences dictated by the examiner. As the subtest progresses, sentence stimuli become longer and linguistically more complex. Scaled scores derived from each subtest will be calculated.

2.9.2.4. Cognitive processing efficiency and executive functions: The Wisconsin Card Sorting Task (WCST-Short Form 64)²⁹ measures strategic planning and cognitive shifting facility, concept-learning efficiency and problem solving. The WCST derives multiple scores that are not designed to be interpreted in isolation. Rather, a holistic approach is taken to provide an overall gestalt of an individual's problem solving strategy. Interpretation of all derived scores provides insight into several factors (e.g. ability to take feedback and alter a problem solving approach; ability to recall a correct strategy; how perseverative a participant may be in approaching this novel task).

The Delis-Kaplin Executive Functioning System (D-KEFS)³⁰ includes: (1) *Trail Making*: this non-verbal measure, akin to an increasingly complex connect-the-dot like task, assesses flexibility-of-thinking and efficiency; (2) *Tower*: this non-verbal block-building task assesses cognitive and motor impulsivity, planning, foresight, and, spatial reasoning, and; (3) *Verbal Fluency*: A game-like format assessing the ability to quickly generate words based on certain categories. These 3 tasks were chosen from the total D-KEFS subtests as representative of significant core domains of foresight/planning, impulsivity/inhibition and abstract reasoning/problem solving. Only the completion time (*Trail Making*), number of correct words (*Verbal Fluency*) and total achievement score and move accuracy ratio (*Tower Test*) are considered.

2.9.2.5. Fine motor development and visual-spatial skills: The Beery Visual-Motor Integration Test (Sixth Ed)³¹ is a straightforward design-copying task for the participant, which begins with simple line-drawings that become increasingly complex. The total number of correctly replicated designs is the variable of interest.

The Grooved Pegboard³² is a timed measure of fine motor speed and dexterity. The participant is asked to insert individual keyed pegs into a pegboard with a single hand. Total completion time, per hand, will be transformed into a standard score and interpreted.

The Beery Visual Perception Test (Sixth Ed)³¹ is a measure of visual discrimination. The participant is asked to simply match increasingly complex geometric designs under timed conditions. The total number of correctly matched designs is the variable of interest.

The Beery Motor Coordination (Sixth Ed)³¹ is a corollary to the Visual Perception and VMI allowing direct comparisons since the instruments are co-normed. The participant is asked to trace designs of increasing complexity again under time constraints. The total number of correctly drawn designs is the variable of interest.

The Judgment of Line Orientation Test³³ is a measure of fine visual discrimination, utilizing 30-items in ascending order of difficulty. The participant is presented a stimulus card of two lines at various angles. They are then shown a multiple choice response card consisting of an array of lines drawn at 18-degree intervals. The participant is required to identify which two lines from the response card match the two lines from the stimulus card. The total correct score is the variable of interest.

2.9.2.6. Language: In the Boston Naming Test³⁴ the participant is asked to name line drawings of various objects ranging from high (e.g. brush) to low (abacus) frequency words and the total number of correctly named objects is tallied.

2.9.2.7. Processing Speed: The Rapid Naming task of the Comprehensive Test of Phonological Processing³⁵ is comprised of two subtests, Rapid Number Naming and Rapid Letter Naming. Rapid Number Naming measures the speed with which an individual can name numbers which are considered an overlearned and easy stimulus. Rapid Letter Naming is a subtest measuring the speed at which an individual can name letters. There are two trials per subtest. Total time in each subtest will be summed and a scaled score will be obtained.

2.9.2.8. Observational Checklists: Observations made by the psychometrists during various portions of testing session (e.g., behavior patterns, oral reading dysfluency) are incorporated. Parents also complete brief checklists while their child is being assessed that provide parent reporting of behavior observed in the home and the neighborhood, including the Achenbach Child Behavior Checklist,³⁶ the Colorado Learning Disabilities Questionnaire,³⁷ and the Behavior Rating Inventory of Executive Function (BRIEF).³⁸

2.9.3 National Center for Toxicological Research-Operant Test Battery (NCTR-OTB)—The NCTR-OTB contains several complex positively-reinforced tasks, in which correct performance is thought to depend on relatively specific and important brain functions which include learning, color and position discrimination, time perception, motivation and short-term memory. In the animal laboratory, aspects of performance of these tasks have been used as metrics for determining the acute and chronic effects of drugs and other chemicals on important aspects of brain function.³⁹ In the clinical setting, the instrument has been used to generate normative data for children as a function of age,⁴⁰⁻⁴³ examine correlations between OTB performance and IQ,⁴⁴ characterize OTB behavioral profiles in different clinical populations (e.g., ADHD, depression, anxiety) as well as assess the ability of therapeutic agents to normalize OTB behaviors in these children,⁴⁵⁻⁴⁷ and generate OTB behaviors for direct comparison with those obtained from laboratory animals, in particular, nonhuman primates.^{45,48-51} Utilization of similar or identical behavioral tasks across species serves not only to aid in the validation of the approach but also to facilitate the interspecies extrapolation of exposure data. Indeed, OTB performance by children is not generally distinguishable from that of well-trained rhesus monkeys. The similarity in OTB performance between monkeys and children^{49,51,52} is of particular importance with regard to extrapolating to humans the neurobehavioral (and possibly neurotoxic) effects of drugs and toxicants as determined in the monkey model. Additionally, the demonstration that several measures of OTB performance correlate highly with measures of intelligence in children^{40-42,44} serves to highlight the relevance of such measures.

The administration of the OTB to children is well-standardized and automated (after a brief introduction by study personnel) utilizing an apparatus that can be set up in any quiet room. The same OTB tasks as used in Paule's study of the effects of ketamine anesthesia on rhesus monkeys are utilized,¹¹ which require approximately one hour.

2.10 Return of results

At the time of consent, parents and adult participants are given the option of obtaining results from the Neuropsychological Assessment Battery. In this battery, each domain is assessed using multiple instruments as described. A neuropsychologist (MJZ) determines what would be accepted as results within the normal range based on his extensive experience with clinical neurodevelopmental testing. For those parents and adult participants who wish to know the results of their testing, if the results for each domain tested are within the normal range, these individuals receive a letter so stating. If the results are not within the normal range, these individuals receive a letter so stating and inviting them to schedule an appointment with the neuropsychologist to review test results and recommendations for management. The Neuropsychological Assessment Battery used in this study is not designed to provide a clinical diagnosis, and this is discussed with the parents.

2.11 Statistical analysis

2.11.1. Sample size/Power considerations—To assess whether exposure to general anesthesia prior to age 3 years is associated with neurodevelopmental abnormalities we will compare several outcomes across two anesthesia exposure groups (single or multiple) with outcomes observed in children not exposed to anesthesia. Two age groups, defined based on the age at which the child is tested (8-12 and 15-19 years), are included. In our previous study,⁵³ we found that the mean group achievement test scores in those with 2 or more exposures to anesthesia were lower than those with no exposure by approximately 0.4 standard deviation units. The scores of those with only 1 exposure to anesthesia were similar to those with no exposure (<0.10 standard deviation difference). The sample size for the present study was determined after weighing both statistical considerations and logistical and resource constraints inherent in the proposed study design. Table 2 presents the statistical power to detect an overall difference across exposure groups using a 2 df test with sample-sizes of N=250, N=150, and N=100 for those with no exposure, 1 exposure and 2 or more exposures for a range of effect sizes consistent with our preliminary data. For pairwise comparisons of multiple exposures and single exposures versus the no exposure group, the proposed sample-sizes will provide statistical power (two-tailed, alpha=0.025) of 80% to detect a difference of 0.37 and 0.32 standard deviation units, respectively.

2.11.2. Analysis—Since a frequency-matched design is used with sex-specific strata defined based on the propensity for receiving general anesthesia prior to age 3 years, the data will be analyzed using mixed models with random effects used to account for the potential clustering of outcomes within sampling strata. Outcomes of interest will include those measures from the neuropsychological assessment battery, as well as the behavioral outcomes from the NCTR-OTB. The Wechsler Abbreviate Scale of Intelligence (WASI) will be of particular interest as an overall measure of intellectual ability, but measures for each of the other domains tested will also be analyzed. For each, anesthesia exposure prior to the age of 3 years (none vs 1 vs 2 or more), age at testing (8-12 years vs 15-19 years) and the age-by-anesthesia interaction effect will be included in the model as explanatory variables. Linear contrasts will be used to compare outcomes between anesthesia exposure groups separately for each age group. In all cases, the least squares mean difference between each anesthesia exposure group versus those with no anesthesia exposure will be

summarized using point estimates and corresponding 95% confidence intervals. Since it is possible that the distributions of some characteristics not available for the calculation of the propensity score may differ between exposure groups, sensitivity analyses will be performed which include these characteristics as additional covariates in the model. In addition, supplemental analyses will also be performed with total duration of general anesthesia included as the explanatory variable, rather than number of exposures. In all cases, distributional assumptions will be assessed with appropriate transformations used as necessary.

Binary outcomes will be analyzed using conditional logistic regression as in our prior studies, taking into account the frequency-matched study design, with results presented using odds ratios and corresponding 95% confidence intervals. The NCTR-OTB data will be quantified as previously described by Paule¹¹ and analyzed using mixed models similar to those described above to take into account the frequency-matched study design.

3. Discussion

3.1. Expected Results

The expected products of this research will be the first detailed phenotype of possible anesthetic-associated neurotoxicity in humans, utilizing a robust population-based patient database and prospectively-administered neuropsychological testing battery. In addition, we will provide the first comparison of effects of anesthetic exposure in children and nonhuman primates utilizing the NCTR-OTB.

Based on our previous observation of cumulative adverse effects of anesthesia on neurocognitive function,^{13,15} we hypothesize that multiple, but not single, exposure(s) to general anesthesia will be associated with impairment in multiple domains of the neuropsychological assessment battery. Given the sparse clinical and pre-clinical literature, it is difficult to predict specifically which domains will be affected. However, based on our work to date and data from nonhuman primates, we speculate that measures of domains related to learning and intellectual ability (e.g., the Wechsler Abbreviated Scale of Intelligence, Comprehensive Test of Phonological Processing) will be affected, whereas those related to memory (e.g. Wide Range Assessment of Memory and Learning) may not be.

Regarding the NCTR-OTB, based on the results of Paule et al,¹¹ we hypothesize that exposure to anesthetics may be associated with impairment of tasks related to learning (e.g., the Incremental Repeated Acquisition task), but not those related to short-term memory (e.g., Delayed Matching-to-Sample). As there were also differences in the effects of anesthesia on acquisition vs. steady-state performance, we also may observe a differential pattern of results at the two age ranges studied (8-12 years vs. 15-19 years). Based on the lack of effect on a behavioral task related to motivation (the Progressive Ratio Task), which may be consistent with our prior finding of a lack of effect of multiple anesthetic exposures on the need for an individual evaluation plan for emotional/behavioral disorders in children,⁹ we would also hypothesize no effect of anesthetic exposure on this task in children. In addition to this opportunity to compare the effects of anesthesia on NCTR-OTB

performance in monkeys and children, we will also have the opportunity to correlate NCTR-OTB results with these standard tests, enhancing our understanding of how the NCTR-OTB applies to human subjects.

3.2. Limitations

Although this study utilizes a large battery of tests, it is not completely comprehensive nor without limitations. All areas of neurodevelopment cannot be assessed in detail in a practicable amount of time. We thus attempt to balance the need for a comprehensive examination of potentially-relevant domains with considerations of feasibility. In some instances (e.g., WRAML-2) we utilize only a subtest of the full instrument, chosen as most representative of the core assessment domains, to minimize testing time.

Outmigration of children born in Olmsted County over the periods of interest but now not available for testing is a potential source of bias which will be evaluated. There is also the potential for participation bias regarding which eligible subjects will agree to participate, with the potential for those parents whose children have been exposed to anesthesia and have learning or other academic challenges being more likely to participate. Because we have access to the medical records (including anesthesia records) of all potential participants, we will be able to compare characteristics of those children who did and did not participate.

The most important limitations of this study are common to all human observational studies of this topic: 1) the inability to distinguish between the effects of anesthesia and the potential effects of the underlying illness, or surgical or diagnostic procedure, that makes anesthesia necessary; and 2) the potential for other unmeasured and unanalyzed confounders that may mediate any observed differences in outcome associated with anesthesia exposure. Given the many factors that may influence outcomes, the latter is a potential problem even with experimental studies. We believe that through our unique population-based approach that can access all available medical records we are in a unique position to measure and account for these potential confounders, but we still cannot exclude this possibility. Thus, like other observational studies, this study will show association, not causality. However, when the evermore compelling pre-clinical results are taken into account, a finding that exposure to anesthesia is associated with a specific phenotype would provide compelling evidence of the relevance of the pre-clinical findings to child health. On the other hand, if no effect is found, it will provide important reassurance that as applied in normal clinical practice, anesthesia does not have long-lasting consequences to child neurodevelopment.

3.3. Implications

The possibility that anesthetics may have lasting detrimental effects on the neurodevelopment of children has profound consequences for clinical care. At present, parents and providers are faced with an accumulating body of both pre-clinical and human data suggesting, but not yet proving, an association between exposure to anesthesia at an early age and a variety of adverse outcomes. The design of the MASK study, using a unique set of resources, aims to provide a detailed phenotype of anesthesia-associated sequelae (if present), as well as important comparative information with phenotypic information in a

nonhuman primate via the NCTR-OTB, that will facilitate the translation of findings from pre-clinical to human studies.

Acknowledgments

This project is supported by grant HD071907 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, and utilizes the resources of the Rochester Epidemiology Project, supported by grant R01-AG034676 from the National Institute on Aging.

We thank the staff at the Psychological Assessment Lab for their assistance in subject testing, and Jake Heier, Summer Undergraduate Research Fellow, for assistance in cohort ascertainment.

This document has been reviewed in accordance with United States Food and Drug Administration (FDA) policy and approved for publication. Approval does not signify that the contents necessarily reflect the position or opinions of the FDA nor does mention of trade names or commercial products constitute endorsement or recommendation for use. The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the FDA.

References

1. Brambrink AM, Evers AS, Avidan MS, et al. Isoflurane-induced neuroapoptosis in the neonatal rhesus macaque brain. *Anesthesiology*. 2010; 112:834–41. [PubMed: 20234312]
2. Olney JW, Ishimaru MJ, Bittigau P, Ikonomidou C. Ethanol-induced apoptotic neurodegeneration in the developing brain. *Apoptosis*. 2000; 5:515–21. [PubMed: 11303910]
3. Olney JW, Wozniak DF, Jevtovic-Todorovic V, Farber NB, Bittigau P, Ikonomidou C. Drug-induced apoptotic neurodegeneration in the developing brain. *Brain Pathol*. 2002; 12:488–98. [PubMed: 12408236]
4. Olney JW, Young C, Wozniak DF, Jevtovic-Todorovic V, Ikonomidou C. Do pediatric drugs cause developing neurons to commit suicide? *Trends Pharmacol Sci*. 2004; 25:135–9. [PubMed: 15019268]
5. Ikonomidou C, Bosch F, Miksa M, et al. Blockade of NMDA receptors and apoptotic neurodegeneration in the developing brain. *Science*. 1999; 283:70–4. [PubMed: 9872743]
6. Fredriksson A, Archer T, Alm H, Gordh T, Eriksson P. Neurofunctional deficits and potentiated apoptosis by neonatal NMDA antagonist administration. *Behav Brain Res*. 2004; 153:367–76. [PubMed: 15265631]
7. Fredriksson A, Ponten E, Gordh T, Eriksson P. Neonatal exposure to a combination of N-methyl-D-aspartate and gamma-aminobutyric acid type A receptor anesthetic agents potentiates apoptotic neurodegeneration and persistent behavioral deficits. *Anesthesiology*. 2007; 107:427–36. [PubMed: 17721245]
8. 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. [PubMed: 12574416]
9. Kodama M, Satoh Y, Otsubo Y, et al. Neonatal desflurane exposure induces more robust neuroapoptosis than do isoflurane and sevoflurane and impairs working memory. *Anesthesiology*. 2011; 115:979–91. [PubMed: 21956042]
10. Satomoto M, Satoh Y, Terui K, et al. Neonatal exposure to sevoflurane induces abnormal social behaviors and deficits in fear conditioning in mice. *Anesthesiology*. 2009; 110:628–37. [PubMed: 19212262]
11. 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. [PubMed: 21241795]
12. Rabbitts JA, Groenewald CB, Moriarty JP, Flick R. Epidemiology of ambulatory anesthesia for children in the United States: 2006 and 1996. *Anesth Analg*. 2010; 111:1011–5. [PubMed: 20802051]

13. Flick RP, Katusic SK, Colligan RC, et al. Cognitive and behavioral outcomes after early exposure to anesthesia and surgery. *Pediatrics*. 2011; 128:e1053–61. [PubMed: 21969289]
14. 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. [PubMed: 22305025]
15. 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. [PubMed: 19293700]
16. 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. [PubMed: 21415431]
17. Hansen TG, Pedersen JK, Henneberg SW, et al. Academic performance in adolescence after inguinal hernia repair in infancy: a nationwide cohort study. *Anesthesiology*. 2011; 114:1076–85. [PubMed: 21368654]
18. 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. [PubMed: 22908104]
19. United States Census Bureau. [Accessed May 1,2014] 2014. at <http://www.census.gov/>
20. Rocca WA, Yawn BP, St Sauver JL, Grossardt BR, Melton LJ 3rd. History of the Rochester Epidemiology Project: half a century of medical records linkage in a US population. *Mayo Clin Proc*. 2012; 87:1202–13. [PubMed: 23199802]
21. Katusic SK, Colligan RC, Barbaresi WJ, Schaid DJ, Jacobsen SJ. Potential influence of migration bias in birth cohort studies. *Mayo Clin Proc*. 1998; 73:1053–61. [PubMed: 9818038]
22. Hopkins, J. Johns Hopkins ACG Case Mix System Version 7.0. Baltimore, MD: Johns Hopkins University; 2005.
23. Baron, I. Neuropsychological evaluation of the child. New York: Oxford University Press; 2004.
24. Strauss, E.; Sherman, EMS.; Spreen, O. A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary. New York: Oxford University Press; 2006.
25. Yeates, KO.; Ris, MD.; Taylor, HG., editors. Pediatric neuropsychology: Research, theory, and practice. New York: Guilford Press; 2000.
26. Wechsler, D. Wechsler Abbreviated Scale of Intelligence. San Antonio, TX: The Psychological Corporation; 1999.
27. Sheslow, D.; Adams, W. Administration and Technical Manual. Second. Wilmington, DE: Wide Range; 2003. Wide Range Assessment of Memory and Learning.
28. Conners, KC. Conners Continuous Performance Test. 2nd. Toronto: Multi Health Systems; 2004.
29. Heaton, RK.; Chelune, GJ.; Talley, JL.; Kay, GG.; Curtiss, G. Wisconsin Card Sorting Test Manual: Revised and Expanded. Odessa, FL: Psychological Assessment Resources; 1993.
30. Delis, DC.; Kaplan, E.; Kramer, JH. Delis-Kaplan Executive Function System: Examiner's manual. San Antonio: The Psychological Corporation; 2001.
31. Beery, K.; Beery, N. The Beery-Buktenica Developmental Test of Visual Motor Integration. Sixth. Bloomington, MN: Pearson; 2010.
32. Matthews, CG.; Klove, R. Instruction Manual for the Adult Neuropsychology Test Battery. Madison, WI: University of Wisconsin Medical School; 1964.
33. Benton AL, Varney NR, Hamsher KD. Visuospatial judgment. A clinical test *Arch Neurol*. 1978; 35:364–7.
34. Kaplan, E.; Goodglass, H.; Weintraub, S. Boston Naming Test. Philadelphia: Lea & Febiger; 1983.
35. Wagner, RK.; Torgesen, JK.; Rashotte, CA. Comprehensive test of phonological processing. Austin, TX: PRO-ED; 1999.
36. Achenbach, TM.; Rescorla, LA. Manual for the ASEBA School-Age Forms & Profiles: An Integrated System of Mult-Informant Assessment. Burlington: University of Vermont, Research Center for Children, Youth & Families; 2001.
37. Willcutt EG, Boada R, Riddle MW, Chhabildas N, DeFries JC, Pennington BF. Colorado Learning Difficulties Questionnaire: validation of a parent-report screening measure. *Psychol Assess*. 2011; 23:778–91. [PubMed: 21574721]
38. Gioia, GA.; Isquith, PK.; Guy, SC.; Kenworthy, L. Behavior Rating Inventory of Executive Function. Odessa, FL: Psychological Assessment Resources; 2000.

39. Paule MG. Chronic drug exposures during development in nonhuman primates: models of brain dysfunction in humans. *Front Biosci.* 2005; 10:2240–9. [PubMed: 15970490]
40. Baldwin RL, Chelonis JJ, Prunty PK, Paule MG. The use of an incremental repeated acquisition task to assess learning in children. *Behav Processes.* 2012; 91:103–14. [PubMed: 22710096]
41. Chelonis JJ, Daniels-Shaw JL, Blake DJ, Paule MG. Developmental aspects of delayed matching-to-sample task performance in children. *Neurotoxicol Teratol.* 2000; 22:683–94. [PubMed: 11106861]
42. Chelonis JJ, Flake RA, Baldwin RL, Blake DJ, Paule MG. Developmental aspects of timing behavior in children. *Neurotoxicol Teratol.* 2004; 26:461–76. [PubMed: 15113607]
43. Chelonis JJ, Gravelin CR, Paule MG. Assessing motivation in children using a progressive ratio task. *Behav Processes.* 2011; 87:203–9. [PubMed: 21507343]
44. Paule MG, Chelonis JJ, Buffalo EA, Blake DJ, Casey PH. Operant test battery performance in children: correlation with IQ. *Neurotoxicol Teratol.* 1999; 21:223–30. [PubMed: 10386825]
45. Baldwin RL, Chelonis JJ, Flake RA, et al. Effect of methylphenidate on time perception in children with attention-deficit/hyperactivity disorder. *Exp Clin Psychopharmacol.* 2004; 12:57–64. [PubMed: 14769100]
46. Chelonis JJ, Edwards MC, Schulz EG, et al. Stimulant medication improves recognition memory in children diagnosed with attention-deficit/hyperactivity disorder. *Exp Clin Psychopharmacol.* 2002; 10:400–7. [PubMed: 12498337]
47. Chelonis JJ, Johnson TA, Ferguson SA, et al. Effect of methylphenidate on motivation in children with attention-deficit/hyperactivity disorder. *Exp Clin Psychopharmacol.* 2011; 19:145–53. [PubMed: 21463072]
48. Mayorga AJ, Popke EJ, Fogle CM, Paule MG. Similar effects of amphetamine and methylphenidate on the performance of complex operant tasks in rats. *Behav Brain Res.* 2000; 109:59–68. [PubMed: 10699658]
49. Paule MG, Forrester TM, Maher MA, Cranmer JM, Allen RR. Monkey versus human performance in the NCTR Operant Test Battery. *Neurotoxicol Teratol.* 1990; 12:503–7. [PubMed: 2247039]
50. Rodriguez JS, Morris SM, Hotchkiss CE, et al. The effects of chronic methylphenidate administration on operant test battery performance in juvenile rhesus monkeys. *Neurotoxicol Teratol.* 2010; 32:142–51. [PubMed: 19737611]
51. Chelonis JJ, Cox AR, Karr MJ, Prunty PK, Baldwin RL, Paule MG. Comparison of delayed matching-to-sample performance in monkeys and children. *Behav Processes.* 2014; 103:261–8. [PubMed: 24440984]
52. Paule MG, Cranmer JM, Wilkins JD, Stern HP, Hoffman EL. Quantitation of complex brain function in children: preliminary evaluation using a nonhuman primate behavioral test battery. *Neurotoxicology.* 1988; 9:367–78. [PubMed: 3143955]
53. Stratmann G, Sall JW, May LD, et al. Isoflurane differentially affects neurogenesis and long-term neurocognitive function in 60-day-old and 7-day-old rats. *Anesthesiology.* 2009; 110:834–48. [PubMed: 19293705]

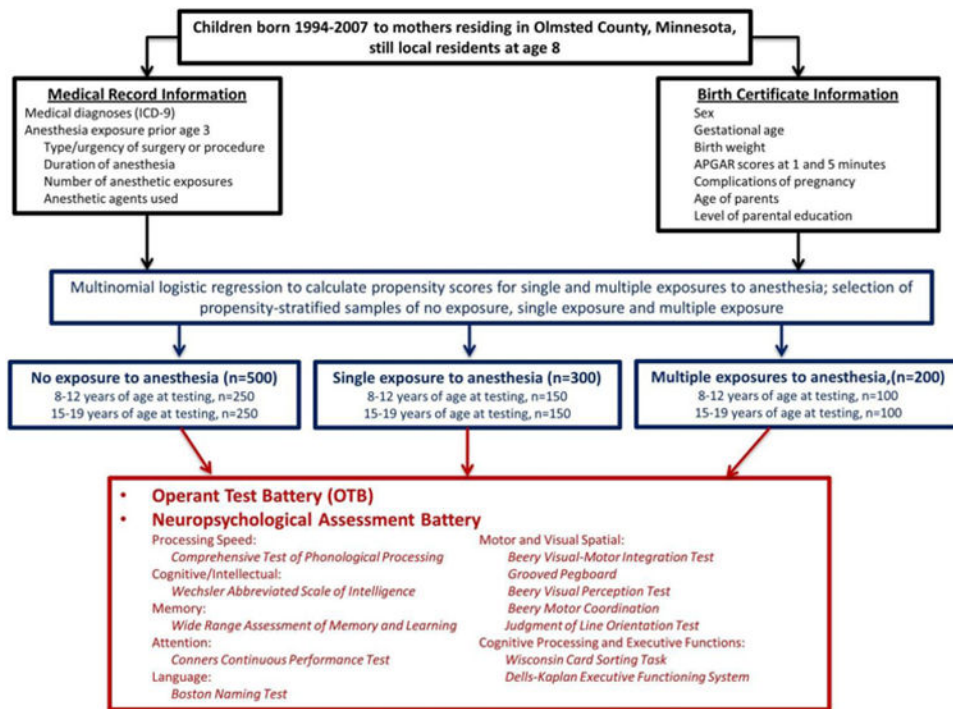


Figure 1. Flow chart describing study data and procedures. Additional details are provided in the Methods section

Table 1
Neuropsychological Assessment Battery testing domains and assessments

Domain	Neuropsychological Test	Specific Assessment
Cognitive	<i>Wechsler Abbreviated Scale of Intelligence (WASI)</i> ²⁶ a. Vocabulary subtest b. Matrix reasoning subtest	Verbal, non-verbal and general intellectual abilities
Memory	<i>Wide Range Assessment of Memory and Learning-2nd Edition (WRAML-2)</i> ²⁷ a. Story-memory task b. Verbal learning task c. Design memory task	Learning/memory (acquisition, storage, retrieval), immediate recall for just-presented material and retention/recall of verbal or visual information
Attention	<i>Conners Continuous Performance Test (2nd Ed)</i> ²⁸ <i>Wide Range Assessment of Memory and Learning-2nd Edition (WRAML-2)</i> ²⁷ a. Finger Windows task b. Number Letter task c. Sentence Memory task	Ability to sustain attention, inhibit incorrect responses, hold information briefly in mind and manipulate it
Executive functioning	<i>Wisconsin Card Sorting Task (WCST-Short Form 64)</i> ²⁹	Problem solving strategy
	<i>Delis-Kaplin Executive Functioning System (D-KEFS)</i> ³⁰ a. Trail making task b. Tower task c. Verbal fluency task	Planning, impulsivity/inhibition and abstract reasoning/problem solving
Fine motor development and visual-spatial skills	<i>Beery Visual-Motor Integration Test (Sixth Ed)</i> ³¹	Fine motor speed and dexterity, visual interpretation and discrimination
	<i>Grooved Pegboard</i> ³²	
	<i>Beery Visual Perception Test (Sixth Ed)</i> ³¹	
	<i>Beery Motor Coordination (Sixth Ed)</i> ³¹	
	<i>Judgment of Line Orientation Test</i> ³³	
Language	<i>Boston Naming Test</i> ³⁴	Confrontational word retrieval
Processing speed	<i>Comprehensive Test of Phonological Processing</i> ³⁵ a. Rapid number naming subtest b. Rapid letter naming subtest	Reading-related phonological processing skills
Multiple brain functions	<i>National Center for Toxicological Research-Operant Test Battery (NCTR-OTB)</i> ⁴⁴	Learning, color and position discrimination, time perception, motivation and short-term memory; utilized for direct comparison between humans and laboratory animals, in particular, nonhuman primates

Table 2
Statistical power for a range of hypothesized effect sizes*

No exposure N=250	One exposure N=150	2 or more exposures N=100	Statistical Power**
0.00 (reference)	0.00	-0.50	98%
0.00	0.00	-0.45	95%
0.00	0.00	-0.40	90%
0.00	0.00	-0.35	81%
0.00	-0.10	-0.50	97%
0.00	-0.10	-0.45	93%
0.00	-0.10	-0.40	86%
0.00	-0.10	-0.35	76%

* Effect sizes for those with one exposure and 2 or more exposures are expressed in standard deviation units with the no exposure group as the reference.

** Statistical power for 2 degree of freedom test using alpha=0.05 to compare anesthesia exposure groups using one-way ANOVA.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript