Title

Standards for preclinical research and publications in developmental anaesthetic neurotoxicity: expert opinion statement from the SmartTots preclinical working group

Journal

British Journal of Anaesthesia, 124(5)

ISSN

0007-0912

Authors

Chinn, Gregory A
Pearn, Matthew L
Vutskits, Laszlo
et al.

Publication Date

2020-05-01

DOI

10.1016/j.bja.2020.01.011

Peer reviewed

BJA

British Journal of Anaesthesia, 124 (5): 585–593 (2020)

doi: 10.1016/j.bja.2020.01.011
Advance Access Publication Date: 4 March 2020
Review Article

NEUROSCIENCE AND NEUROANAESTHESIA

Standards for preclinical research and publications in developmental anaesthetic neurotoxicity: expert opinion statement from the SmartTots preclinical working group

Gregory A. Chinn[1], Matthew L. Pearn[2], Laszlo Vutskits[3], Cyrus D. Mintz[4], Andreas W. Loepke[5], Jennifer J. Lee[6], Jerri Chen[6], Zeljko J. Bosnjak[7], Ansgar M. Brambrink[6], Vesna Jevtovic-Todorovic[8], Lena S. Sun[6] and Jeffrey W. Sall[1,*]

[1]Department of Anesthesia and Perioperative Care, University of California, San Francisco, CA, USA, [2]Department of Anesthesiology, University of California, San Diego, CA, USA, [3]Department of Anesthesiology, Clinical Pharmacology, Intensive Care and Emergency Medicine, University Hospitals of Geneva, Geneva, Switzerland, [4]Department of Anesthesiology and Critical Care Medicine, The Johns Hopkins University School of Medicine, Baltimore, MD, USA, [5]Department of Anesthesiology and Critical Care Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA, USA, [6]Department of Anesthesiology, Columbia University, New York, NY, USA, [7]Department of Medicine, Medical College of Wisconsin, Milwaukee, WI, USA and [8]Department of Anesthesiology, University of Colorado, Aurora, CO, USA

*Corresponding author. E-mail: Jeffrey.Sall@ucsf.edu

Summary

In March 2019, SmartTots, a public–private partnership between the US Food and Drug Administration and the International Anesthesia Research Society, hosted a meeting attended by research experts, anaesthesia journal editors, and government agency representatives to discuss the continued need for rigorous preclinical research and the importance of establishing reporting standards for the field of anaesthetic perinatal neurotoxicity. This group affirmed the importance of preclinical research in the field, and welcomed novel and mechanistic approaches to answer some of the field's largest questions. The attendees concluded that summarising the benefits and disadvantages of specific model systems, and providing guidance for reporting results, would be helpful for designing new experiments and interpreting results across laboratories. This expert opinion report is a summary of these discussions, and includes a focused review of current animal models and reporting standards for the field of perinatal anaesthetic neurotoxicity. This will serve as a practical guide and road map for novel and rigorous experimental work.

Keywords: anaesthesia; animal model; neurodevelopment; neurotoxicity; paediatric anaesthesia; research guidelines; research reporting

Editor's key points

- Preclinical research has been instrumental in identifying and characterising anaesthetic neurotoxicity.
- An expert panel was convened to recommend standards for future research and reporting in this field.
- These recommendations include a focused review of current animal models and reporting standards for the field of perinatal anaesthetic neurotoxicity.

Received: 26 June 2019 **Accepted:** 24 January 2020

The SmartTots consortium is a public–private partnership between the US Food and Drug Administration (FDA) and the International Anesthesia Research Society (IARS). Its mission is to facilitate research that ensures safe anaesthesia and sedation for millions of children who undergo surgery or procedures that require anaesthesia each year. In March 2019, a group of preclinical research experts in this field, editors from leading anaesthesia journals, and representatives from the FDA and funding agencies met with the support of SmartTots to discuss the need for high-quality preclinical studies with consistency in reporting methods and results. In the nearly 20 yr since the alarm was raised in neonatal rats,[1] preclinical investigations have consistently shown neuroanatomical changes and lifelong cognitive deficits after exposure to common anaesthetics in early life.[2,3] This has sparked clinical studies, both prospective and retrospective, that have corroborated the consistent preclinical finding that single anaesthetic exposures of short duration in otherwise healthy individuals do not cause long-term deficits, but concerns from both preclinical and clinical studies regarding longer duration and multiple exposures remain unsettled.

As with any other scientific question, different model systems present unique advantages and limitations, and therefore, investigations of developmental anaesthetic neurotoxicity have been conducted using a wide range of organisms and systems. *In vitro* models as varied acute brain slice cultures to human embryonic stem cells provide mechanistic and physiological insights to important but narrow questions. Animal models offer the best practical approach to testing potential strategies to minimise or mitigate risks with a goal of identifying a non-toxic anaesthesia technique.[4] Studies in animals are critical to understanding the phenotype resulting from developmental anaesthetic neurotoxicity, which can be difficult to discern in human patients, in which the confounding influences of surgery, co-morbid disease, socio-economic and environmental factors, etc. are difficult to control. Although progress has been made in this arena, further research in animal models is essential to understanding the mechanism(s) of transient and permanent anaesthetic effects on the brain, as there is no feasible approach to conducting mechanistic studies in humans. Studies in animals must be carefully designed, rigorously executed, and thoroughly reported to achieve this goal.

A number of different approaches have demonstrated anaesthetic neurotoxicity in a wide range of organisms and systems. Although confirmatory, these different approaches present challenges when comparing results. In addition, the maturation of the field and technological developments have increased the threshold of novelty that merits publication in high-quality journals. The aim of this expert opinion report was to provide guidance to improve novelty, rigour, and reproducibility for preclinical research of anaesthetic neurotoxicity by establishing commonly agreed-upon standards and to support meaningful advances in research design and implementation.

We propose a framework for approaching preclinical questions regarding the effects of anaesthetics on neuronal development. Appropriate design and reporting standards for preclinical studies will enable comparative analyses across studies and correlation with available human data. Such carefully conducted, reliably reported, and reproducible anaesthetic-mediated developmental neurotoxicity research may expand our knowledge of fundamental processes important to brain development and neuronal plasticity, and to mechanisms of anaesthetic action. This report provides an analysis with recommendations to guide and challenge investigators in this field to follow and adhere to best practices in preclinical research, experimental design, and model system selection.

Methods

This article is a synthesis of discussions that occurred at the preclinical SmartTots meeting held from April 31 to May 1, 2019 in New York City. Attendees included basic science researchers with active research interests in this area, journal editors, and members of the FDA and IARS (see online Supplementary material). The specifics of the process were discussed before adjournment on the last meeting day and were agreed upon by all present. Briefly, participants were divided into four groups of eight to 10 people with group leaders who are all experts in the field. The group topics were *animal models*, *research methods*, *publication and reporting*, and *future directions*. The group leaders were tasked with leading and summarising the small group discussions, presenting the small group summary to the entire group, and providing written notes of the entire group and conference organisers after the meeting. This was synthesised into a manuscript by several junior attendees, and circulated and edited by the group leaders and conference organisers. All conference participants, representatives from the FDA, and the SmartTots advisory board were given an opportunity to provide input after the initial draft was complete. Final suggestions were incorporated and the completed manuscript submitted for publication.

Models

The choice of preclinical model of neonatal anaesthesia exposure depends on the particular question to be answered. There are multiple different experimental models in the literature, each with relative advantages and disadvantages. Here, we highlight the most common preclinical models, and discuss the advantages and inherent limitations for their use in this type of research (Table 1).

In vitro models

As long as fundamental questions regarding anaesthetic mechanism(s) and cellular toxicity remain, there will continue to be a role for *in vitro* models. These models include established cell lines, primary cell cultures, acute and organotypic slice cultures, human stem cell cultures, and three-dimensional organoids. Each model can be exquisitely controlled and molecular mechanisms and electrophysiology can be interrogated at precise time points in single-cell types or reductionist systems. Connectivity can be explored depending on the complexity of existing circuits. They also allow for inexpensive, high-throughput screening experiments, and could prove valuable in the development of novel anaesthetics or protective agents. The use of *in vitro* models is preferred over the use of live animals, given increasing ethical concerns for using animals in research.[5,6]

Such reductionist models have certain disadvantages, but when used appropriately and in concert with *in vivo* models, scientists can leverage their simplicity for great potential. Nonetheless, it is important to recognise the limitations of *in vitro* models relative to *in vivo* systems, as even the best *in vitro* models consist of damaged tissue and disrupted

Table 1 Comparison of animal models used in preclinical anaesthetic neurotoxicity research.

Animal model	Advantages	Disadvantages
Cell culture	Economy Least ethical concerns	No or limited connectivity Translation unclear
Brain slice culture	Focus on cellular mechanisms	Lack of long-term outcomes
3D organoids	Insights into molecular and genetic mechanisms	Lack complexity, including immune and vascular systems that are essential for maturity and function
Non-mammalian species:	Space-efficient housing	Maturational stages unclear
Nematode (*Caenorhabditis elegans*)	Genetic and molecular mechanisms	No cognitive outcomes
Fruit fly (*Drosophila melanogaster*)	apply to humans	No physiological monitoring
Zebrafish (*Danio rerio*)	Suitable for toxic screens Mechanistic studies Genetic manipulations	
Small rodents	Smallest animal models with interacting neuronal networks Established, validated cognitive tests Equivalent tests available in humans (startle reflex) Genetic manipulations Prenatal human brain stages can be modelled postnatally	Limitations in physiological monitoring Brain morphology dissimilar to humans (lissencephaly)
Guinea pigs		Precocious brain development compared with humans (*in utero* comparable with term human neonates)
Pigs	Gyrencephalic brain similar to humans Most economic large animal models Full physiological monitoring available	Limited cognitive assessment
Sheep		Precocious brain development compared with humans (*in utero* comparable with term human neonates) Relatively high cost
Non-human primates	Closest to human morphology and physiology	Ethical considerations Limited supply High cost

circuits in artificial media under non-homeostatic conditions. Direct comparison of cell culture assays with *in vivo* development is inherently flawed given the artificial environment that *in vitro* systems necessitate. Although it is possible to correlate specific developmental events that occur in cells *in vitro* with analogous events for the same cell type *in vivo*, the timescales between the two models are not necessarily the same. Rather than drawing tenuous connections between *in vitro* and *in vivo* systems, it may be better to acknowledge the limitations of the model and define the parameters meticulously. These models should be supported by measuring expression of developmental molecular markers and physiological variables. Rather than generalising findings of *in vitro* work to animal models or even clinical scenarios, it may be more powerful to focus on mechanistic actions that can be interrogated in these simplified systems.

In vivo models

These models cover a range of species from flies and worms to non-human primates (NHPs). One of the distinct advantages of animal models over *in vitro* models is the ability to interrogate drug effects on animal behaviour. Here, we highlight the most commonly used animal models and identify appropriate areas of use.

Non-mammalian species

Several non-mammalian species are routinely used as model systems in the investigation of anaesthetic effects on the developing nervous system, including *Caenorhabditis elegans*, zebrafish, and *Drosophila melanogaster*. Space-efficient housing, animal costs, robust genetic tools, and short life cycles make these models useful for toxicological screening, mechanistic studies, and exploration of the interactions between genes and anaesthetics. What remain unclear are the equivalent developmental states of these organisms in comparison with mammalian species. In addition, fewer cognitive assays are available for use in non-mammalian species and monitoring of physiological factors is difficult. Finally, it is a challenge to compare exposure levels in non-mammalian species with plasma exposures in humans.

Mammalian species

The majority of model systems that have been used are mammalian, with a heavy reliance on small rodents. Here, we highlight some of the advantages and disadvantages of small rodents that are specific to this field, and explore the role for larger mammalian models.

One challenge for any animal model that translates brain development to humans has been the inherent differences in maturation between species. Computational models, such as those published on translatingtime.org,[7,8] offer estimates of equivalent post-conceptional dates across mammalian species using empirical neural events for comparison. This can be useful for testing hypotheses regarding specific developmental events; however, it should be recognised that computational approaches necessarily

simplify the complex process of brain development that limit its extrapolation.

Small rodents

Rodent models are the most commonly used mammalian models for studying anaesthetic neurotoxicity. However, these models have been criticised, in particular regarding both the lack of physiological control and monitoring of physiological parameters during anaesthetic exposure.[9–11] However, because rodents possess a brain cytoarchitectonic organisation comparable with humans, they remain a powerful model to study neonatal anaesthetic toxicity and the cornerstone of paediatric drug development programmes. The tremendous growth in the field of murine genetic engineering has made mice particularly useful for identifying molecular mechanisms involved in neuronal development and toxicity. Batteries of established assessments of behaviour, anxiety, and memory are available for mice and rats.[12,13] These rodent behavioural tests have corollary methods that can be applied in primates and humans,[14–17] allowing for potential translation from rodent to human studies should a relevant phenotype within a subpopulation of rodents be identified. Another advantage is that the rodent brain develops substantially in early postnatal life; this allows for an easy study of events that occur *in utero* and early neonatal life in humans. The genetic tools available for mice make them ideal for certain types of mechanistic and proof-of-concept experiments. Rats may be preferred to other rodents with regard to their consistency in behavioural tasks, which can result in fewer animals needed per experimental group.[18–20]

Guinea pigs

Guinea pigs possess some of the advantages and disadvantages of mice and rats. In contrast to smaller rodents, brain development in guinea pigs is more precocious, meaning that their third trimester *in utero* development mirrors the early postnatal period in humans.[21] The longer length of gestation in guinea pigs compared with mice (59–72 *vs* 20 days, respectively),[22] can allow the study of anaesthetic exposure duration relative to the rate of brain development. As the larger pregnant guinea pig allows appropriate physiological monitoring, this model provides the opportunity of maintaining physiological homeostasis.[23] However, this approach precludes the randomisation of littermates to different exposure groups, and the lack of genetic tools and species-specific reagents can be a significant drawback that contributes to the paucity of studies using this animal model. Guinea pigs require more space relative to smaller rodents even when socially housed, which can incur increased costs.

Pigs

Porcine models have substantial advantages over rodent models, mostly related to size and neuroanatomy. The larger size allows for accurate physiological monitoring, anaesthetic delivery via tracheal intubation and mechanical ventilation, and ease of venous and arterial vascular access.[24–26] The gyrencephalic brains of pigs closely resemble humans in both appearance and developmental sequence.[27] Behavioural tests include corollaries to many commonly tested domains in rodents and NHPs.[28] However, swine behavioural models are less established compared with rodents because of fewer published studies and the different testable behaviours in pigs *vs* rodents.[28] Imaging studies might be particularly fruitful in this model as their larger size means that human scanners (CT or MRI) can be utilised.[29] This is advantageous over smaller animals whose brains are particularly difficult to image and require specialised scanners with extremely powerful magnets.[30] Furthermore, the timing of the rapid increase in brain size, known as the *brain spurt*, is most closely aligned between pigs and humans compared with the other mammals reviewed here.[21] The disadvantages of pigs include the costs of animals and husbandry, and the lack of species-specific biochemical reagents, such as antibodies, gene sequences, and primers. However, they are significantly less expensive and have a much shorter gestation period relative to NHPs (115 *vs* 165 days in pigs *vs* NHPs, respectively).[31,32]

Sheep

The maturation stages of neonatal sheep are precocious with the end of the third trimester corresponding to early postnatal human development.[33] Despite differences in developmental time, this gyrencephalic model has been used in other fields of paediatric neurology to model vascular disruptions and their effects on brain development.[34] Like guinea pigs and swine, the physiological parameters in this large mammalian model can be accurately monitored and controlled. However, the *in utero* exposure required for brain developmental equivalency to human neonates precludes randomisation to different exposure groups for littermates of the same gestation. Similar to other large animal species, the limitations on behavioural assessments and lack of species-specific reagents, and the relative cost, pose significant challenges to the use of sheep models.[35]

Non-human primates

Non-human primate research offers perhaps the strongest preclinical evidence for the human relevance of developmental anaesthetic neurotoxicity observed in animals.[36–45] The comparable physiology, parallel neuronal development, and complex behaviours all contribute to the importance of this model. NHP models offer the closest evolutionary model to human brain development. However, this model should be limited to the most clinically relevant questions to maximise the scientific gain from their limited use.[46,47] NHPs remain critical for corroborating experimental findings in lower mammals. The primary limitations are the ethical concerns, costs, and time required for these studies, and a much more limited set of tools to investigate cellular and molecular mechanisms. Although critically important for linking findings in lower vertebrates to humans, NHP studies cannot be relied on to guide this field.

Reporting standards

With the maturation of this field of research, there is an increasing need for rigorous and thoughtfully designed experiments disseminated with clear reporting standards. This is crucial for the replication of experiments and for the comparison and interpretation of results across different research groups, models, and experimental designs. Here, we present the best practice guidelines for performing experiments and reporting methods and results in developmental anaesthetic neurotoxicity.

General considerations

Improving the design of preclinical studies with regard to rigour and reproducibility has been the subject of ongoing debate for the past decade in research. In fact, the US National Institutes of Health (NIH) has published the Principles and Guidelines for Reporting Preclinical Research[48] that mirrors recommendations from a prominent white paper.[49] Similarly, the UK National Centre for the Replacement, Refinement & Reduction of Animals in Research published the Animal Research: Reporting of In Vivo Experiments guidelines for reporting preclinical animal studies commonly referred to as the ARRIVE guidelins.[6] These have been widely adopted as best practice by journals, although they are rarely enforced in their entirety for all published studies. The document, as a checklist, is freely available.[50] Although detailed and structured, the specifics do not apply universally. However, the authors of this report strongly recommend adhering to these standards. The following are recommendations specific to the field of anaesthetic toxicity.

Anaesthesia exposure

In addition to the basics of agent, dose, and length of exposure, details regarding the delivery method, depth of anaesthesia, and monitoring of the agent should be clearly stated. For inhaled gases, the relative concentration should be directly measured, not solely relying on delivery settings. Anaesthetic records should be kept with frequent measurements of volatile concentrations and other physiological parameters. The volumetric composition of the carrier gas should be reported for both the anaesthesia and control groups. Mortality and adverse events must be reported. Although the cause of death may be unknown, speculation can be helpful for interpretation of results and adjustment of methods to mimic clinical relevance.

For experiments aiming to define a therapeutic index and determine safety margins, *in vivo* studies should also be designed to include multiple doses and exposure durations that might define adverse effect exposures, in which no adverse effects are observed (no-observed-adverse-effect level). When feasible, obtaining toxicokinetic data can be extremely helpful to extrapolate exposures across species.

Physiological data

Various species present different opportunities to monitor physiological parameters during anaesthetic exposure. Temperature remains a minimum essential variable to report when studying all neonatal animals. Temperature measured directly by skin or other method (not only ambient temperature) should be recorded several times an hour, particularly in the smallest animals. For animals larger than rodents, blood pressure in neonates can be reliably taken either by a non-invasive blood pressure cuff or by arterial cannulation.[25,51] Although there are reports of successful pulse oximetry monitoring in neonatal rodents,[52] practical limitations in animal size and device accuracy for neonatal rodents have limited widespread use. Accurate pulse oximetry is readily achievable in larger models and should be used.[25,53]

End-tidal CO_2 should be monitored and recorded for models of volatile anaesthetic exposure when delivered by tracheal intubation, and can be measured in anaesthetic chambers for smaller animals. Care should be taken to limit inhaled CO_2 as much as possible, and can include using CO_2 absorbers. In larger animals, which readily allow tracheal intubation, methods for ventilation should be recorded along with the technical description of the intubation procedure.

In the absence of continuous monitoring of oxygenation and ventilation, arterial blood gas analysis is critical for establishing a new model system, or for new investigators to the field attempting to replicate what other laboratories have reported. Each laboratory should report the impact of the anaesthetic protocol on arterial blood gases at least once. In an effort to reduce rodent use, we advise measuring arterial blood gases only when a group is publishing an initial study from their laboratory, or if there have been changes made in the anaesthetic protocol (e.g. regarding the duration of exposure or dose).

Nutritional status

Given the time of pups away from their mothers during anaesthesia exposure, it is possible for nutritional status to affect outcomes. Intermittent recording of weight should be done after the anaesthetic exposure, especially for new models of anaesthesia exposure, validation of a new method, or survival studies. Litter size, particularly if culled, should be noted, as this can influence nutritional status.

Sex as a biological variable

Animals of both sexes should be included in a study if appropriate, and the distribution of male/female subjects should be reported.[54] It should be specified whether sex was analysed as a biological variable. The US NIH[55] recommends that, 'sex as a biological variable will be factored into research designs, analyses, and reporting in vertebrate animal and human studies. Strong justification from the scientific literature, preliminary data, or other relevant considerations must be provided for applications proposing to study only one sex'. Within the field of anaesthetic neurotoxicity, sex has become an important variable with differences in behavioural and neurochemical responses to anaesthetics between males and females.[56–58]

Experimental sample size

Ideally, sample size should be determined by *a priori* power analysis.[59] However, effect size is often unknown before performing the experiments. Furthermore, the number critically depends on the relevant primary outcome, physiologically and pharmacologically significant differences, number of comparisons, type of data, and statistical testing. In rodent behavioural studies, groups of 12–20 are commonly required to detect differences between two groups with an effect size of 1.0–1.5 at an alpha of 0.05 and power of 0.95. In contrast, immunohistochemistry effect sizes are often as large as two to three, so groups of three to six animals may be appropriate to maintain the same alpha and power. The FDA, the US Environmental Protection Agency, and the Organization for Economic Cooperation and Development generally recommend a sample of 20 animals per sex per group for developmental neurotoxicity studies.[60] In the case of large animals, in which individual animals are particularly expensive, robust phenotypes should be prioritised to utilise appropriate cohort sizes and limit underpowered studies.

An important criticism of animal work in general has been that studies are often highly underpowered.[61–63] This leads to problems with interpretation and decreases reproducibility.[63] It also further complicates broader interpretation of outcomes within the field, leads to delays, and requires additional studies to arrive at agreed-upon conclusions. Although there is a continuing push to reduce and replace the number of animals used in biomedical research, this should not come at the expense of adequate statistical power. Appropriate powering of a study by using more animals could ultimately reduce total animal use, as underpowered studies that lead to erroneous or ambiguous results might need to be repeated, or lead to unnecessary studies to explore unsubstantiated results.

Animal husbandry

Animal husbandry, particularly in the perinatal period, plays an important role in animal development and subsequent behaviour.[28,64–67] Therefore, care should be taken to minimise neonatal and maternal stressors during this period. The sources of animals (i.e. vendor *vs* breeding) and day (age) of arrival should be described. The type of cage (i.e. with or without enrichment or ventilation) should also be described, as this can influence subsequent behaviour. Details regarding weaning (age), separation of animals, and housing number should also be reported.

Post-anaesthesia animal care should be described (i.e. what criteria were used to return animals to their mothers). Special care, including rubbing animals with bedding, may help cue rodent dams to continue to care for infants and may prevent rejection.[68] Light/dark cycles and time of day of experiments should be noted, as sleep and learning/behavioural performance are critically linked, and some behaviours are more sensitive to performance during resting or active periods.[69–71]

The order of behavioural assessments should be considered and reported. Details, such as the sequence of behavioural assessments, days of rest, time of day, and sex of the test administrator, can significantly influence animal behaviour results.[72,73] For multiple or repeated studies, care should be taken to replicate testing conditions as closely as possible for all animals and all groups.

Experimental design/outcomes

Experiments should be transparently reported in detail. If detailed reporting of a method has been published, it is acceptable to reference an earlier study for details. Established methods usually have best practices. For example, with immunoblots, a sample size of more than two biological replicates is preferred, in addition to technical replicates that originate from the same animal,[74,75] and the entire blot must be available for critical evaluation.[75] In primary cell culture experiments, at least two separate cultures should be investigated for true replication.[76] A best practice approach should apply to other methods commonly used, such as immunofluorescence, electron microscopy, and electrophysiology, and should include the number of animals the sections originated from, how many sections per animal were evaluated, and how the sections evaluated were chosen. For all studies, the sample size (n) should be clearly defined, such that data can be interpreted easily and other investigators can replicate the technique and compare results. Methods for blinding, randomisation, and sampling should be clearly described.

Behavioural experiments should be sufficiently detailed for replication and comparison with other studies. Reporting should include the protocol, administrators of testing, time of the day, and details of the environment where testing occurred. Exclusion of individual animals during testing should be reported and justified in detail. Positive and negative controls should be carefully considered especially if conclusions are made regarding negative outcomes.

Statistical analysis

Power analyses should ideally be conducted *a priori* using effect size from similar experiments, even if this may be difficult for preclinical pilot studies. Statistical approaches for analysing data should be clearly reported in a specific statistical method section. Data sets should be made available for independent analysis, in particular for large data studies, such as genetic screens. The discussion of excluded data should be explicit with appropriate rationale.

Negative data and controls

Bias against publishing negative results continues in science in general.[77,78] We encourage investigators to publish negative results with transparent experimental methods with appropriate positive controls critical for the interpretation of these observations. Negative controls are also critical for verifying the accuracy of assays. Editors and reviewers are encouraged to consider well-designed and conducted studies on important topics for publication, even if their results are negative.

Redundant methods

The conclusions drawn from experiments with redundant methods are inherently stronger given the testing of the hypothesis by different means. There is a need for rigorously testing hypotheses by multiple methods if possible, especially for negative data. An example is validating experiments using both a receptor knockout model and a receptor-blocking drug.

Raw data

For small data sets, individual data points should be displayed on graphs so the interpretation of scatter and outliers can be assessed.[79] Summarising data with bars can be misleading and hide the effects of outliers, and should be used with caution or when sample size makes display of all points difficult. In accordance with the push to make science more open, we recommend that raw data be submitted to journals when the option is available to allow for an independent analysis.[80,81] If data hosting is not possible, we encourage providing data to investigators upon request.

Ethical treatment of animals

Animal studies must undergo an institutional review. A statement that includes the approved protocol and attestation to the ethical treatment of animals should be included.

Conflict of interest and funding source

Authors must state funding sources and list potential conflicts of interest including financial (Industry and government contracts or travel and consultanting fees) and personal interests (stock holdings and company ownership).

Conclusions

The ultimate goal of preclinical research is to enhance biomedical discovery by conducting experiments that are difficult or impossible in humans, but may lead to improved clinical care and understanding. As with other clinically related preclinical research, laboratory investigations in the field of developmental anaesthesia neurotoxicity face practical questions of scaling animal work to humans, such as equating age of exposure or exposure length in animals to brain maturation and exposure times in humans, comparing i.v. with i.p. injections, and extrapolating the significance of observed behavioural outcomes. This area of research is particularly challenging, given the absence of a simple robust phenotype in humans. Expected outcomes from neurotoxic exposure are cognitive deficits that require complex neuropsychological testing to be detected. Nevertheless, the SmartTots workshop participants agree that the current clinical data show signals in behavioural domains and secondary outcomes that may be related to findings observed in preclinical work. Future research is necessary to define these domains in both healthy and potentially vulnerable (i.e. with co-morbidities) populations of children who require sedation or anaesthesia.

Results from experimental research will continue to expand our understanding of anaesthetic neurotoxicity and can provide guidance on how to design future human studies. Whether such investigations involve genetic mechanisms of toxicity in nematodes or rodents, imaging and connectivity in large mammals, or carefully titrating doses to establish toxicity levels in NHP, these experiments will be valuable and informative to the study of the effects of anaesthetics in particular and developmental neuroscience in general. Rigorous experimental design and reporting will provide a higher standard in advancing the field toward the collective goal of improving anaesthetic care and outcomes in young children.

Authors' contributions

Synthesis of discussions from the meeting: GAC, MLP
Article structure/scope conceptualisation: LSS, JWS
Writing initial draft: GAC, MLP
Article editing: all authors.

Declaration of Interest

LV is an editor of *Anesthesiology*. VJT is a member of the associate editorial board of the *British Journal of Anaesthesia*. The other authors declare that they have no conflicts of interest.

Acknowledgements

The meeting was supported by grant 5U01FD005935-03 from the US Food and Drug Administration.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bja.2020.01.011.

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. Lin EP, Lee JR, Lee CS, Deng M, Loepke AW. Do anesthetics harm the developing human brain? An integrative analysis of animal and human studies. *Neurotoxicol Teratol* 2017; **60**: 117–28
3. Vutskits L, Xie Z. Lasting impact of general anaesthesia on the brain: mechanisms and relevance. *Nat Rev Neurosci* 2016; **17**: 705–17
4. Soriano SG, Vutskits L, Jevtovic-Todorovic V, Hemmings HC. BJA Neurotoxicology and Neuroplasticity Study Group. Thinking, fast and slow: highlights from the 2016 BJA seminar on anaesthetic neurotoxicity and neuroplasticity. 2017 *Br J Anaesth* 2016; **119**: 443–7
5. Russell W, Burch R. *The Principles of humane experimental technique*. London: Methuen; 1959
6. Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. *PLoS Biol* 2010; **8**, e1000412
7. Workman AD, Charvet CJ, Clancy B, Darlington RB, Finlay BL. Modeling transformations of neurodevelopmental sequences across mammalian species. *J Neurosci* 2013; **33**: 7368–83
8. Finlay BL, Darlington RB. Linked regularities in the development and evolution of mammalian brains. *Science* 1995; **268**: 1578–84
9. Stratmann G, May LD, Sall JW, et al. Effect of hypercarbia and isoflurane on brain cell death and neurocognitive dysfunction in 7-day-old rats. *Anesthesiology* 2009; **110**: 849–61
10. Wu B, Yu Z, You S, et al. Physiological disturbance may contribute to neurodegeneration induced by isoflurane or sevoflurane in 14 day old rats. *PLoS One* 2014; **9**, e84622
11. Johnson SC, Pan A, Sun GX, et al. Relevance of experimental paradigms of anesthesia induced neurotoxicity in the mouse. *PLoS One* 2019; **14**, e0213543
12. Carter M, Shieh J. Animal behavior. In: Carter M, Shieh J, editors. *Guide to research techniques in neuroscience*. 2nd Edn. San Diego, CA: Academic Press; 2015. p. 39–71
13. Hånell A, Marklund N. Structured evaluation of rodent behavioral tests used in drug discovery research. *Front Behav Neurosci* 2014; **8**: 252
14. Manns JR, Hopkins RO, Reed JM, Kitchener EG, Squire LR. Recognition memory and the human hippocampus. *Neuron* 2003; **37**: 171–80
15. Burgess N, Maguire EA, O'Keefe J. The human hippocampus and spatial and episodic memory. *Neuron* 2002; **35**: 625–41
16. Shrager Y, Bayley PJ, Bontempi B, Hopkins RO, Squire LR. Spatial memory and the human hippocampus. *Proc Natl Acad Sci U S A* 2007; **104**: 2961–6
17. Schoenfeld R, Schiffelholz T, Beyer C, Leplow B, Foreman N. Variants of the Morris water maze task to comparatively assess human and rodent place navigation. *Neurobiol Learn Mem* 2017; **139**: 117–27
18. Frick KM, Stillner ET, Berger-Sweeney J. Mice are not little rats: species differences in a one-day water maze task. *Neuroreport* 2000; **9**: 3461–5

19. Whishaw IQ, Tomie J-A. Of mice and mazes: similarities between mice and rats on dry land but not water mazes. *Physiol Behav* 1996; **60**: 1191–7
20. Stranahan AM. Similarities and differences in spatial learning and object recognition between young male C57Bl/6J mice and Sprague-Dawley rats. *Behav Neurosci* 2011; **125**: 791–5
21. Dobbing J, Sands J. Comparative aspects of the brain growth spurt. *Early Hum Dev* 1979; **3**: 79–83
22. Anon. The all-purpose guinea pig. *Lab Anim* 2014; **43**: 79
23. Rizzi S, Carter LB, Ori C, Jevtovic-Todorovic V. Clinical anesthesia causes permanent damage to the fetal guinea pig brain. *Brain Pathol* 2008; **18**: 198–210
24. Whitaker EE, Zheng CZ, Bissonnette B, et al. Use of a piglet model for the study of anesthetic-induced developmental neurotoxicity (AIDN): a translational neuroscience approach. *J Vis Exp* 2017: 55193
25. Whitaker EE, Bissonnette B, Miller AD, et al. A novel, clinically relevant use of a piglet model to study the effects of anesthetics on the developing brain. *Clin Transl Med* 2016; **5**: 2
26. Broad KD, Kawano G, Fierens I, et al. Surgery increases cell death and induces changes in gene expression compared with anesthesia alone in the developing piglet brain. *PLoS One* 2017; **12**, e0173413
27. Lind NM, Moustgaard A, Jelsing J, Vajta G, Cumming P, Hansen AK. The use of pigs in neuroscience: modeling brain disorders. *Neurosci Biobehav Rev* 2007; **31**: 728–51
28. Gieling ET, Nordquist RE, van der Staay FJ. Assessing learning and memory in pigs. *Anim Cogn* 2011; **14**: 151–73
29. Sauleau P, Lapouble E, Val-Laillet D, Malbert C-H. The pig model in brain imaging and neurosurgery. *Animal* 2009; **3**: 1138–51
30. Denic A, Macura SI, Mishra P, Gamez JD, Rodriguez M, Pirko I. MRI in rodent models of brain disorders. *Neurotherapeutics* 2011; **8**: 3–18
31. Garnett I, Rahnefeld GW. Factors affecting gestation length in the pig. *Can J Anim Sci* 1979; **59**: 83–7
32. Silk J, Short J, Roberts J, Kusnitz J. Gestation length in rhesus macaques (*Macaca mulatta*). *Int J Primatol* 1993; **14**: 95–104
33. McIntosh GH, Baghurst KI, Potter BJ, Hetzel BS. Foetal brain development in the sheep. *Neuropathol Appl Neurobiol* 1979; **5**: 103–14
34. Back SA, Riddle A, Hohimer AR. The sheep as a model of brain injury in the premature infant. In: Yager JY, editor. *Animal models of neurodevelopmental disorders*. New York, NY: Springer; 2015. p. 107–28
35. Morton AJ, Avanzo L. Executive decision-making in the domestic sheep. *PLoS One* 2011; **6**, e15752
36. Raper J, De Biasio JC, Murphy KL, Alvarado MC, Baxter MG. Persistent alteration in behavioural reactivity to a mild social stressor in rhesus monkeys repeatedly exposed to sevoflurane in infancy. *Br J Anaesth* 2018; **120**: 761–7
37. Raper J, Alvarado MC, Murphy KL, Baxter MG. Multiple anesthetic exposure in infant monkeys alters emotional reactivity to an acute stressor. *Anesthesiology* 2015; **123**: 1084–92
38. Alvarado MC, Murphy KL, Baxter MG. Visual recognition memory is impaired in rhesus monkeys repeatedly exposed to sevoflurane in infancy. *Br J Anaesth* 2017; **119**: 517–23
39. Liu F, Rainosek SW, Frisch-Daiello JL, et al. Potential adverse effects of prolonged sevoflurane exposure on developing monkey brain: from abnormal lipid metabolism to neuronal damage. *Toxicol Sci* 2015; **147**: 562–72
40. Schenning KJ, Noguchi KK, Martin LD, et al. Isoflurane exposure leads to apoptosis of neurons and oligodendrocytes in 20- and 40-day old rhesus macaques. *Neurotoxicol Teratol* 2017; **60**: 63–8
41. Creeley CE, Dikranian KT, Dissen GA, Back SA, Olney JW, Brambrink AM. Isoflurane-induced apoptosis of neurons and oligodendrocytes in the fetal rhesus macaque brain. *Anesthesiology* 2014; **120**: 626–38
42. Creeley C, Dikranian K, Dissen G, Martin L, Olney J, Brambrink A. Propofol-induced apoptosis of neurones and oligodendrocytes in fetal and neonatal rhesus macaque brain. *Br J Anaesth* 2013; **110**: i29–38
43. Coleman K, Robertson ND, Dissen GA, et al. Isoflurane anesthesia has long-term consequences on motor and behavioral development in infant rhesus macaques. *Anesthesiology* 2017; **126**: 74–84
44. Brambrink AM, Back SA, Riddle A, et al. Isoflurane-induced apoptosis of oligodendrocytes in the neonatal primate brain. *Ann Neurol* 2012; **72**: 525–35
45. Brambrink AM, Evers AS, Avidan MS, et al. Isoflurane-induced neuroapoptosis in the neonatal rhesus macaque brain. *Anesthesiology* 2010; **112**: 834–41
46. Phillips KA, Bales KL, Capitanio JP, et al. Why primate models matter. *Am J Primatol* 2014; **76**: 801–27
47. Anon. When less is not more. *Nat Med* 2008; **14**: 791
48. National Institutes of Health. *Principles and guidelines for reporting preclinical research* 2014. Available from: http://www.nih.gov/about/reporting-preclinical-research.htm. [Accessed 20 February 2020]
49. Landis SC, Amara SG, Asadullah K, et al. A call for transparent reporting to optimize the predictive value of preclinical research. *Nature* 2012; **490**: 187–91
50. *National Centre for the replacement refinement & reduction of animals in research. ARRIVE guidelines checklist*. Available from: https://www.nc3rs.org.uk/arrive-guidelines. [Accessed 20 February 2020]
51. Giguère S, Knowles Jr HA, Valverde A, Bucki E, Young L. Accuracy of indirect measurement of blood pressure in neonatal foals. *J Vet Intern Med* 2005; **19**: 571–6
52. Takada SH, Sampaio CAG, Allemandi W, Ito PH, Takase LF, Nogueira MI. A modified rat model of neonatal anoxia: development and evaluation by pulseoximetry, arterial gasometry and Fos immunoreactivity. *J Neurosci Methods* 2011; **198**: 62–9
53. Dawson JA, Bastrenta P, Cavigioli F, et al. The precision and accuracy of Nellcor and Masimo oximeters at low oxygen saturations (70%) in newborn lambs. *Arch Dis Child Fetal Neonatal* 2014; **99**: F278–81
54. Cabrera OH, Gulvezan T, Symmes B, Quillinan N, Jevtovic-Todorovic V. Sex differences in neurodevelopmental abnormalities caused by early life anaesthesia exposure: a review. Br J Anaesth Adv.Access Published on January 21, 2020, doi: 10.1016/j.bja.2019.12.032.
55. National Institutes of Health. *Consideration of sex as a biological variable in NIH-funded research (NOT-OD-15-102)*. Available from: https://orwh.od.nih.gov/sites/orwh/files/docs/NOT-OD-15-102_Guidance.pdf. [Accessed 20 February 2020]. Release date 9 Jun 2015
56. Lee BH, Chan JT, Kraeva E, Peterson K, Sall JW. Isoflurane exposure in newborn rats induces long-term cognitive

dysfunction in males but not females. *Neuropharmacology* 2014; **83**: 9–17
57. Chastain-Potts SE, Tesic V, Tat QL, Cabrera OH, Quillinan N, Jevtovic-Todorovic V. Sevoflurane exposure results in sex-specific transgenerational upregulation of target IEGs in the subiculum. *Mol Neurobiol* 2020; **57**: 11–22
58. Ju L-S, Yang J-J, Morey TE, et al. Role of epigenetic mechanisms in transmitting the effects of neonatal sevoflurane exposure to the next generation of male, but not female, rats. *Br J Anaesth* 2018; **121**: 406–16
59. Fitts DA. Ethics and animal numbers: informal analyses, uncertain sample sizes, inefficient replications, and type I errors. *J Am Assoc Lab Anim Sci* 2011; **50**: 445–53
60. Organisation for Economic Cooperation and Development. *Test no. 426: developmental neurotoxicity study* 2017. Available from: https://www.oecd-ilibrary.org/environment/test-no-426-developmental-neurotoxicity-study_9789264067394-en. [Accessed 20 February 2020]
61. Higginson AD, Munafò MR. Current incentives for scientists lead to underpowered studies with erroneous conclusions. *PLoS Biol* 2016; **14**, e2000995
62. Munafò MR, Nosek BA, Bishop DVM, et al. A manifesto for reproducible science. *Nat Hum Behav* 2017; **1**: 21
63. Button KS, Ioannidis JPA, Mokrysz C, et al. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* 2013; **14**: 365
64. Weinstock M. Prenatal stressors in rodents: effects on behavior. *Neurobiol Stress* 2016; **6**: 3–13
65. Beery AK, Kaufer D. Stress, social behavior, and resilience: insights from rodents. *Neurobiol Stress* 2014; **1**: 116–27
66. Turner KM, Burne THJ. Comprehensive behavioural analysis of Long Evans and Sprague-Dawley rats reveals differential effects of housing conditions on tests relevant to neuropsychiatric disorders. *PLoS One* 2014; **9**, e93411
67. Hsu F-C, Zhang G-J, Raol YSH, Valentino RJ, Coulter DA, Brooks-Kayal AR. Repeated neonatal handling with maternal separation permanently alters hippocampal $GABA_A$ receptors and behavioral stress responses. *Proc Natl Acad Sci U S A* 2003; **100**: 12213–8
68. Rincón-Cortés M, Sullivan RM. Emergence of social behavior deficit, blunted corticolimbic activity and adult depression-like behavior in a rodent model of maternal maltreatment. *Transl Psychiatry* 2016; **6**: e930
69. Fisk AS, Tam SKE, Brown LA, Vyazovskiy VV, Bannerman DM, Peirson SN. Light and cognition: roles for circadian rhythms, sleep, and arousal. *Front Neurol* 2018; **9**: 56
70. Orts-Sebastian A, Ludin NM, Pawley MDM, Cheeseman JF, Warman GR. Impact of anaesthesia on circadian rhythms and implications for laboratory experiments. *Exp Neurol* 2019; **311**: 318–22
71. Sara SJ. Sleep to remember. *J Neurosci* 2017; **37**: 457–63
72. Gulinello M, Mitchell HA, Chang Q, et al. Rigor and reproducibility in rodent behavioral research. *Neurobiol Learn Mem* 2019; **165**: 106780
73. Kafkafi N, Agassi J, Chesler EJ, et al. Reproducibility and replicability of rodent phenotyping in preclinical studies. *Neurosci Biobehav Rev* 2018; **87**: 218–32
74. Fosang AJ, Colbran RJ. Transparency is the key to quality. *J Biol Chem* 2015; **290**: 29692–4
75. Brooks HL, Lindsey ML. Guidelines for authors and reviewers on antibody use in physiology studies. *Am J Physiol Heart Circ Physiol* 2018; **314**: H724–32
76. Lazic SE, Clarke-Williams CJ, Munafò MR. What exactly is 'N' in cell culture and animal experiments? *PLoS Biol* 2018; **16**, e2005282
77. Dalton JE, Bolen SD, Mascha EJ. Publication bias: the elephant in the review. *Anesth Analg* 2016; **123**: 812–3
78. Anon. The importance of no evidence. *Nat Hum Behav* 2019; **3**: 197
79. Weissgerber TL, Milic NM, Winham SJ, Garovic VD. Beyond bar and line graphs: time for a new data presentation paradigm. *PLoS Biol* 2015; **13**, e1002128
80. National Academy of Sciences. *Open science by design: Realizing a vision for 21st century research*. Washington, DC: The National Academies Press; 2018
81. Royal Society. *Science as an open enterprise*. London: Royal Society; 2012

Handling editor: Hugh C Hemmings Jr