SUPPLEMENTAL GLOSSARY

Clinical studies which set out to understand causation use a specific and often technical vocabulary, which can prove challenging for a non-expert audience. In order to facilitate the reading of our main paper, we present a glossary defining the key terms used. Much of the vocabulary used in our paper comes from the field of epidemiology: the “study of the distribution and determinants of health-related states or events” (1). In epidemiological studies, researchers try to identify the causal pathways that lead from different factors to health outcomes or diseases. These factors and outcomes are often defined as “variables”, with analyses conducted to identify the relationship between them. Most epidemiological research is by necessity observational – relying on observed data from individuals and populations, rather than that derived from planned experiments or trials. This is because in many situations it is impossible, or unethical, to modify the factors that might cause a health outcome or disease. An example of this are the potential neurodevelopmental consequences of general anaesthetic agents in young children.

In animals, the neurotoxic effects of anaesthetics have been demonstrated in rodent (2) and non-human primate models (3–6), and there is thus the possibility that they might also lead to detrimental effects in the human developing brain. However, for several surgical procedures it is clearly unethical not to provide general anaesthesia, and thus it is not possible to compare the effects of anaesthetics in one group of children, with the absence of the use of these in another control group. Thus, research looking at the effects of prolonged anaesthesia has generally focused on observational data comparing infants and children exposed to different degrees and types of anaesthesia.
In the glossary below, we outline the terms used to define variables in causal thinking, with reference to the example of the potential adverse consequences of anaesthetic agents. As with all research into causality, it will become clear that a major preoccupation in the field is trying to identify true causal relationships, whilst avoiding confusing these with non-causal associations between variables, or failing to identify causal relationships when they are in fact present. As a reference for compiling this glossary we have drawn extensively on the CDC publication “Principles of Epidemiology in Public Health Practice” (7).

**BIAS**

In causal studies, one wants to understand the true relationship between the putative cause of a health outcome (see Exposure, below) and that outcome (see Outcome, below). In the case of anaesthetics and neuro-development, one would like to know whether anaesthetic agents do in fact harm the developing brain, whether certain anaesthetics are more likely to do so than others, and the characteristics of the dose response relationship. A bias is any process at any stage of inference which tends to produce results or conclusions that differ systematically from the truth (8), and affects the validity of the study.

Bias can therefore be used as a general term to indicate systematic – as opposed to random – distortion of inference. This broad, philosophical definition can be made more specific by realising that some threats to validity are due to conditions that exist in nature (that the researcher must disentangle and attempt to control for), which amount to “confounding”, and some that do not exist in nature, but are created by the researcher by his/her decisions in design, inclusion criteria, measurement, follow-up, and analysis, which we refer to in this glossary and throughout our review as “bias” proper.
In this example, the researcher might reach incorrect (“biased”) conclusions for several reasons. The exposed and unexposed groups might differ in other features in addition to the exposure— for instance, children undergoing surgery and anaesthesia could be sicker, and hence at greater neurodevelopmental risk, than controls (Confounding, see below). Equally, bias might arise from the ways in which data are measured, for example if the assessors who ascertained outcome were not blinded to exposure (Information bias, see below); or from the selective loss to follow-up of some patients in the compared groups, for instance, due to less active follow-up of children operated on with an unremarkable clinical course (Selection bias, see below). Bias differs from random error in that it is systematic, that is, not due to chance, and a sample size increase will not obviate it. In this glossary and throughout our review we will maintain the convention that bias is a feature of the study, rather than the causal relationships being investigated, therefore differentiating bias from confounding.

There are however several classifications of bias, with substantial variation in nomenclature (9), that differ across authors and disciplines. For instance, the Cochrane collaboration labels as “attrition bias” what we define here as “selection bias”, and calls “detection bias” what we define as “information bias”. The most significant difference is in what the Cochrane collaboration calls “selection bias”, which they define as a bias due, for instance, to faulty random allocation resulting in a baseline imbalance (10). In this review, as most often in the field of epidemiology, however, this is generally called “confounding”. In the social sciences, bias is often defined more generally as a threat to validity, and in particular to internal validity: that is, that the causal relationships described are in fact a correct representation of the biological reality.
CONFOUNDING

Confounding is a mixing of effects, where “the apparent effect of the exposure of interest is distorted, because the effect of extraneous factors is mistaken for – or mixed with - the actual exposure effect (which may be null)”(11). Older definitions reported that a confounding variable is a factor that is related to both an exposure (see below) and a health outcome even among those not exposed, but is not in the causal pathway between the two. Confounding exists when a confounding variable influences the outcome of interest, and is differentially distributed between those exposed and those not exposed, leading to an imbalance between the two groups at baseline. For instance, several studies (although not all) (12) have shown a relationship between prolonged or repeated exposure to anaesthesia and poorer neuro-cognitive outcomes (13–16). However, it is possible that the strength of this relationship could be distorted by a confounding variable: the children undergoing more procedures could have more significant underlying medical conditions, such as congenital anomalies, which in themselves confer a higher risk of neuro-developmental impairment (12). That is, it is not the general anaesthetic that causes, or is the only cause of the adverse neuro-developmental outcomes, but other factors associated with having a general anaesthetic such as the underlying diagnosis.

More generally, in analysing the consequences of medical or surgical interventions, it is often difficult to disentangle the effect of the intervention from the reason for which it was administered. For example, it is difficult to ascertain the effect of caesarean section (CS) compared with vaginal delivery on the survival of extremely preterm infants, because women undergoing CS differ from those not undergoing it for several key features which can influence outcome, such as the type of pregnancy complication leading to delivery, socio-
economic status, antenatal care, or steroid prophylaxis. This is often called “confounding by indication”. Where feasible, randomisation ensures (on average) that this mechanism does not bias results, and represents the great advantage of randomised clinical studies over observational studies. It is important to note that – except when randomisation successfully removes these unwanted associations – real-world situations always present a pattern of confounders, and these will lead to a threat to the validity of the study unless they are adequately accounted for in their design or analysis. To reach correct conclusions, the task of the researcher is therefore to correct for confounding.

EXPOSURE

In epidemiology, an exposure is a variable that is thought to be a possible cause of a health outcome or disease, and which is chosen as the focus of investigation by the researcher. In our example, the exposure is the administration of anaesthetic agents, and is treated as variable that can be related to other variables (such as the health outcome, or potential confounders) in subsequent analyses. The exposure of interest can be classified differently, and thus it is important to examine how a study has defined an exposure, and how this has been measured, before being able to compare the results from different studies. In the case of anaesthesia and neurodevelopment, the type of anaesthetic agents used has changed over time (14). In addition, some studies categorise exposure to anaesthesia on the basis of diagnostic codes for a surgical procedure (14,17), whereas others classify exposure by the number of hours of general anaesthetic (18).

INFORMATION BIAS
Information bias is where there is a systematic difference, between the groups being compared, in the way that information about the exposure and outcome is collected, leading to differential misclassification. For example, the outcome measures might not be measured in standardised way, and could thus be systematically affected by lack of blinding and the assessor’s knowledge of which group an individual was allocated to. One can imagine that the neurodevelopment of a child with complex medical conditions undergoing surgical procedures would be assessed differently to that of an otherwise healthy child. If there are systematic biases in the way that an exposure or outcome is measured in different groups, this could lead to a distortion of the true relationship between the two. Some authors have suggested that the discrepancy in the results of different studies looking into neurodevelopment after anaesthesia could be due to the diversity of outcome measures used both within, and between, studies (18).

A type of information bias is recall bias, which can occur in a case-control study where the information offered on the exposure differs between cases and controls (differential misclassification). An example of this is a study which asked parents of children with acute lymphoblastic leukaemia, and matched controls, the distance from their residence to the nearest power line (19). This was in the context of a population concerned about a perceived excess of leukaemia cases (which they related to proximity to electricity cables). The researchers found that the parents of those with leukaemia, compared to controls, tended to underestimate the distance from their residence to the nearest power cable. This meant that the odds ratio for the association between leukaemia and the distance to the nearest power line was inflated when using parental recall of this distance, rather than an objective measure: differential misclassification had introduced bias into the results.
MEDIATOR

Sometimes also called an intermediate or intermediary variable, a mediator has been defined as “any factor that represents a step in the causal chain between the exposure and disease [that] should not be treated as an extraneous confounding factor, but instead requires special treatment” (11). That is, in the causal pathway, the exposure causes the mediator, which in turn causes the health outcome.

OUTCOME

The outcome is the health-related event caused by the exposure; in our example a child’s neurodevelopment. The outcome variable can again be assessed in variety of ways. In this example, it could be measured in terms of diagnostic codes from medical records documenting learning disability or attention deficit hyperactivity disorder (17), performance in a group administered achievement test (20) or individually administered test (21), documentation in school records (14), or findings on brain imaging (18).

SELECTION BIAS

This is one of the most vexed definitions of all, where significant differences exist between disciplines. As alluded to above, we do not use this term to describe the situation where the groups being studied differ in their baseline characteristics due to selection procedures, and this imbalance could (theoretically) be amended by randomization – we call this confounding. Nor is there selection bias simply when the studied population is not representative of the source population. Randomised controlled trials, where the populations
studied are precisely defined by inclusion and exclusion criteria leading to “selection” from the source population, show that no bias necessarily ensues from this lack of population representativeness, although it may well affect the generalisability of the findings (the external validity).

As described in our review, selection bias ensues when inclusion (selection) in a study differs between groups depending on features associated with the exposure and the outcome. Of note, we use here inclusion in a study in a broad sense, that is not limited to enrolment, but including completion of the study, and collection and analysis of data. That is, the exposed and not-exposed, or subjects with and without the outcome, have a different probability of being studied. Typically, this is due to systematic differences in follow-up rates between the groups being compared. For instance, if children undergoing surgery/anaesthesia are lost to follow-up more frequently when their clinical course is unremarkable, this would lead to a spurious enrichment of cases with bad prognosis, and to an overestimation of the risk due to anaesthesia (differential selection on outcome). We deal with selection bias in more detail in the main body of the review.
References

1. WHO | Epidemiology. WHO 2014;

2. 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.

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

4. 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.

5. 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.

6. Briner A, Nikonenko I, De Roo M, Dayer A, Muller D, Vutskits L. Developmental Stage-dependent Persistent Impact of Propofol Anesthesia on Dendritic Spines in the Rat Medial Prefrontal Cortex. Anesthesiology 2011;115:282–93.

7. Principles of Epidemiology: Glossary|Self-Study Course SS1978|CDC [Internet]. [cited 2018 Mar 9];Available from: https://www.cdc.gov/OPHSS/CSELS/DSEP/SS1978/Glossary.html#B

8. Murphy EA. The logic of medicine. Baltimore: Johns Hopkins UP; 1976.

9. Sackett DL. Bias in analytic research. J Chronic Dis 1979;32:51–63.

10. Higgins JPT, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration’s tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.

11. Rothman K, Greenland S, Lash T. Validity in epidemiologic studies. In: Modern Epidemiology. 2008. p. 128–47.

12. Bartels M, Althoff RR, Boomsma DI. Anesthesia and Cognitive Performance in Children: No Evidence for a Causal Relationship. Twin Res Hum Genet 2009;12:246–53.

13. 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.

14. Flick RP, Katusic SK, Colligan RC, et al. Cognitive and Behavioral Outcomes After Early Exposure to Anesthesia and Surgery. Pediatrics 2011;128:e1053–61.

15. 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.

16. Hu D, Flick RP, Zaccariello MJ, et al. Association between Exposure of Young Children to Procedures Requiring General Anesthesia and Learning and Behavioral Outcomes in a Population-based Birth Cohort. Anesthesiology 2017;127:227–40.

17. DiMaggio C, Sun LS, Kakavouli A, Byrne MW, Li G. A Retrospective Cohort Study of the Association of Anesthesia and Hernia Repair Surgery With Behavioral and
Developmental Disorders in Young Children. J Neurosurg Anesthesiol 2009;21:286–91.

18. Backeljauw B, Holland SK, Altaye M, Loepke AW. Cognition and Brain Structure Following Early Childhood Surgery With Anesthesia. Pediatrics 2015;136:e1-12.

19. Bosco JLF, Silliman RA, Thwin SS, et al. A most stubborn bias: no adjustment method fully resolves confounding by indication in observational studies. J Clin Epidemiol 2010;63:64–74.

20. Hansen TG, Pedersen JK, Henneberg SW, et al. Academic Performance in Adolescence after Inguinal Hernia Repair in Infancy. Anesthesiology 2011;114:1076–85.

21. 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.