BJA

British Journal of Anaesthesia, 125 (3): 393–397 (2020)

doi: 10.1016/j.bja.2020.03.031
Advance Access Publication Date: 27 June 2020
Review Article

STATISTICS IN ANAESTHESIA

Causal inference in perioperative medicine observational research: part 1, a graphical introduction

Vijay Krishnamoorthy[1,*], Danny J. N. Wong[2], Matt Wilson[3], Karthik Raghunathan[1,4], Tetsu Ohnuma[1], Duncan McLean[5], S. Ramani Moonesinghe[3] and Steve K. Harris[3]

[1]Critical Care and Perioperative Epidemiologic Research (CAPER) Unit, Department of Anesthesiology, Duke University Hospital, Durham, NC, USA, [2]Department of Anaesthesia, Guy's and St Thomas' NHS Foundation Trust, London, UK, [3]Critical Care, University College London Hospitals NHS Foundation Trust, London, UK, [4]Anesthesiology Service, Durham Veterans Affairs Medical Center, Durham, NC, USA and [5]Department of Anesthesiology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

*Corresponding author. E-mail: vijay.krishnamoorthy@duke.edu

Abstract

Graphical models have emerged as a tool to map out the interplay between multiple measured and unmeasured variables, and can help strengthen the case for a causal association between exposures and outcomes in observational studies. In Part 1 of this methods series, we will introduce the reader to graphical models for causal inference in perioperative medicine, and set the framework for Part 2 of the series involving advanced methods for causal inference.

Keywords: causal inference; confounding; epidemiology; graphical models; observational research

Editor's key points

- This review provides an introduction to graph theory as applied to perioperative medicine.
- Concepts of causality, bias, and confounding are addressed.
- The biggest threats to validity in observational research are spurious correlations and systematic bias.
- A graph theoretical approach can be useful in mitigating these threats.

The clinician's goal is to directly improve the health of the patient in their care; unfortunately, the research data available to clinicians for decision-making can vary in methodological rigour. Among the myriad of 'associations' between variables reported in the clinical literature, it is imperative for clinicians and researchers to identify true causal relationships. Therefore, we need to understand which exposures are causally connected to better health outcomes.

The validity of medical research can be compromised by two potential sources of error: (1) spurious correlation (or spurious lack of correlation), which are akin to Types 1 and 2 statistical errors, respectively, and (2) systematic biases, which involve biases in the process of the overall study design, lead to spurious interpretation/evaluation, and are non-random in nature. Statistics provides the tools to establish a correlation (or association) between variables, and this is required before delving into causality. Using the language of probability, statistics allows us to distinguish signal from noise. Examining births in London between 1664 and 1757, and finding 737 629 male and 698 958 females (an excess of 38 671 male births), we could conclude that something is amiss.[1] However, scale this back 10 000-fold, and observing 74 male and 70 female births, we might have little to say. The probability of the former occurring by chance is 1.08e−228, while the latter is 0.40 (40%). These probabilities are numerical estimates of plausibility, and the machinery to calculate them developed from investigating games of chance in the 18th century. From this follows classical statistical hypothesis testing, Bayesian techniques, and machine learning.

Received: 22 August 2019; **Accepted:** 17 March 2020

Unfortunately, statistics often does not have the tools to handle systematic bias, which can impair the ability to establish a causal association between variables. For example, we can be confident that people carrying cigarette lighters are likely to have higher rates of lung cancer, but we do not know whether banning cigarette lighters will save lives. Not only does association not imply causation, but it provides no insight into the direction of causality. The association would be identical whether the lung cancer causes a craving for cigarettes, or cigarettes cause lung cancer. Thus, the observed association between two variables may be explained by a causal pathway in either direction or a non-causal association. Systematic biases can result from a conceptual model that has misspecified the relationship of key biological variables, despite the fact that the statistical analyses are sound (mathematical equations are equivalent whether written forwards or backwards). The solution to this problem has been rigorous experimental design, where the process allows us to distinguish cause from effect. Just as we are warned that 'association is not causation', there is a complementary truism that, there is 'no causation without manipulation'.[2]

The grammar for causality abandons algebra and probability for a theory of graphs. The field of graphical models has been developed by Judea Pearl,[3] and won him the Turing Prize (the 'Nobel prize of computing') in 2011. The graph invites us first to add cigarettes as a common cause of carrying lighters and lung cancer. It provides tools to decide whether we can be confident that the observed relationship is free from systematic bias or not. Finally, it offers guidance on the measurements we need to make to determine whether the flow of causality is from the cigarette or the lighter. The joy of this visual approach is its accessibility, although there has been—as yet—little penetration of these methods into the clinical literature.[4]

This review will introduce this language through a worked example developed from a study of long-term survival after major surgery.[5] We will then use this new tool to discuss the causal estimation of treatment effects. This will provide a framework for Part 2 of this series, which introduces the reader to advanced methods for causal inference in perioperative medicine.

Table 1 Commonly used terms in graphical modelling.

Term	Definition
Directed Acyclic Graph	Causal model; no node is visited twice (acyclic)
Node	Variable of interest
Edge	Arrow demonstrating relationship between two variables, indicating a *possible* causal effect between two variables
Children	Downstream variables (at the head of the arrow)
Descendants	Downstream variables beyond 'children'
Parents	Upstream variables (at the tail of the arrow)
Ancestors	Upstream variables beyond 'parents'
Path	Sequence of nodes and edges
Confounder	Variable that is associated with an exposure and outcome, but is not on the causal pathway
Collider	Variable that is influenced by two other (potentially unrelated) variables
Backdoor path	A non-causal relationship between two variables

Graphical causal models

Worked example: long-term survival after major surgery

Khuri and colleagues[5] published a report in 2005 from the Veterans Affairs hospitals in the USA examining the relationship between postoperative complications and long-term survival. This was an observational study, *not* an RCT. Measures of preoperative risk, surgical severity, and postoperative complications were linked with long-term outcome data. Postoperative complications were seen as a proxy for poor perioperative care, and the authors wished to quantify the importance of this period of care. For expository purposes, we will only consider a small subset of the measurements used.

Graphical models: the syntax of a model

Before we sketch the graphical model of this study, we need to introduce the vocabulary. Graphs are composed of 'nodes (variables) and edges (arrows).' The arrow indicates a possible causal effect between two variables. Downstream variables (at the head of the arrow) are often called children; upstream variables (at the tail) are ancestors. The direct ancestors are the 'parents' of a variable.

The absence of an arrow is a statement of a strong assumption that there is no causal effect.[4] Conversely, the presence of an arrow does not mandate that the parent is a cause of the child; the causal effect may be zero. Therefore, where we cannot assume there is no causal relation, we must include an arrow. A path is a sequence of arrows. For causality to be possible, a path cannot be circular (egg → chicken → egg); each node on a path can only be visited once. Where this is true, then the model may be called a Directed Acyclic Graph (DAG). Because of the assumption that the absence of an arrow implies the absence of an effect, then a Causal DAG is a statement of everything that is known about a particular process.

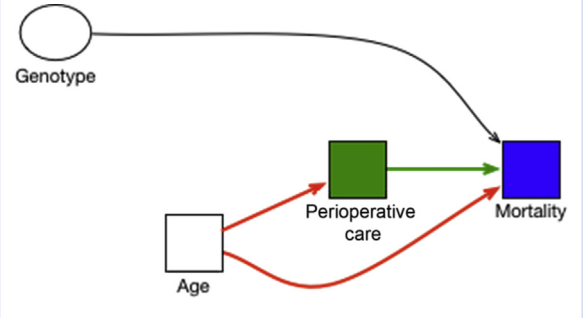

Fig 1. Worked example of graphical causal model highlighting the relationship between the exposure (perioperative care) in green, and the outcome (mortality) in blue. The green arrow highlights the causal relationship we are trying to estimate. The red arrows highlight paths that create non-causal (biasing) connections between the exposure and the outcome.

That completes the formal definition of the components of a graphical model, as discussed in Table 1. However, one will often see additional colouring and shading which is used to aid discussion. Nodes are often drawn as squares to indicate that a variable is observed, or as circles where the variable is unobserved or unmeasured. In our worked example (Fig 1), we highlight the exposure (perioperative care) in green, and the outcome (mortality) in blue. We use a green arrow to highlight the causal relationship we are trying to estimate (the effect of perioperative care on long term outcomes). Finally, we may use red to highlight paths that create non-causal connections between the exposure and the outcome.

The use of graphical models

Each model starts with the two nodes, the exposure and the outcome, linked by one edge: the causal relationship of interest. Remember that the arrow is only indicating a possible effect; this may be zero in the final analysis. We then add all other possible relationships to the graph, allowing us to decide if we can identify the causal relationship. This is only possible if we can eliminate all non-causal paths between the exposure and the outcome. This means finding a set of measures that controls for confounding, without introducing selection bias through a collider. We will now expand our model of perioperative care and mortality to illustrate this process.

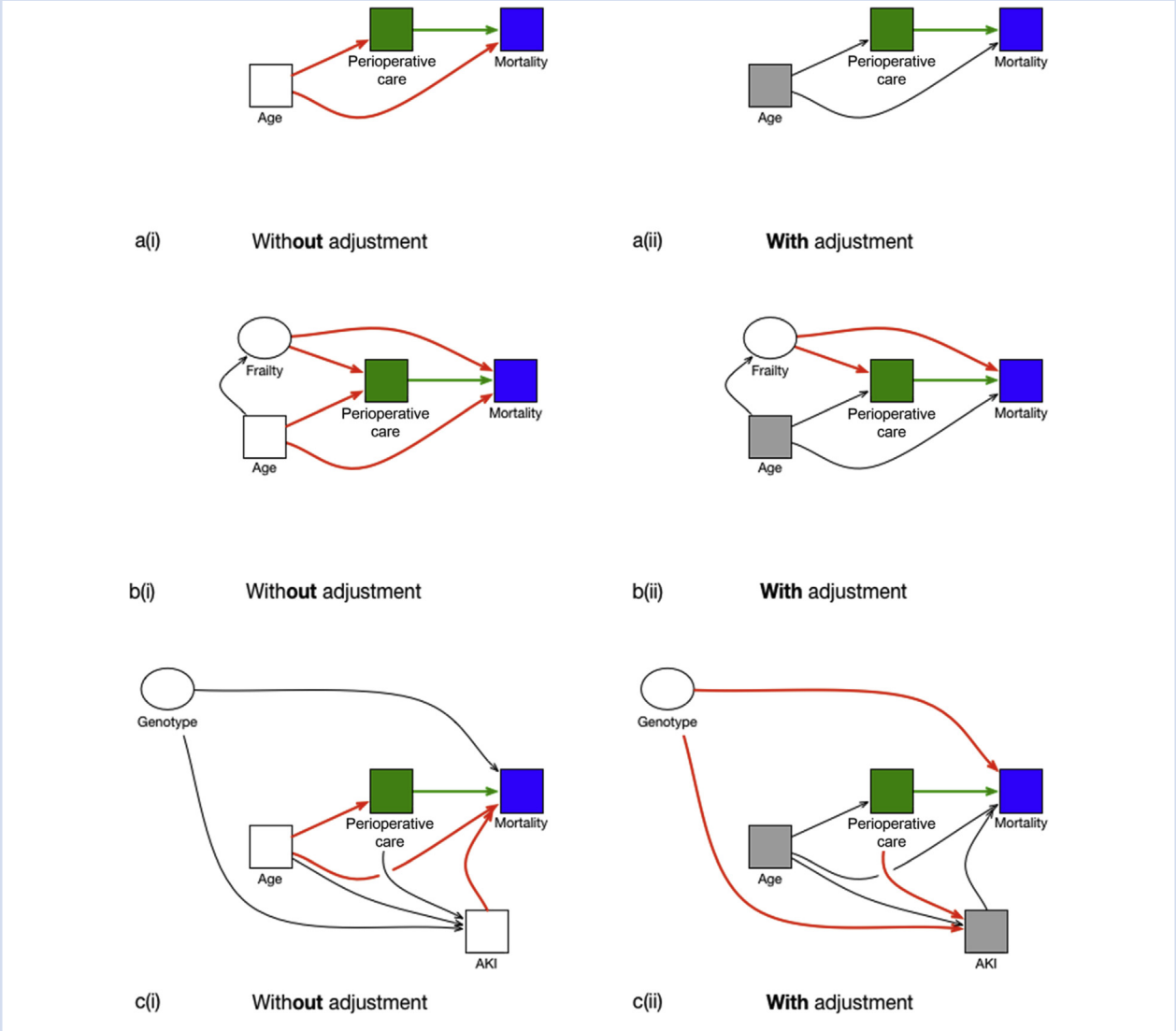

Fig 2. Worked examples of graphical causal model highlighting the relationship between the exposure (perioperative care) in green, and the outcome (mortality) in blue. These examples illustrate the relationships that confounding influences have on exposures and outcomes, and how adjusting for confounders can mitigate this. For example, (c, i)–>(c, ii) is a situation where uncritically adjusting for potential confounders has inadvertently opened up a backdoor/biasing path through a variable which is unmeasured and therefore cannot be adjusted for, resulting in bias. AKI, acute kidney injury.

Confounding and backdoor paths

Two variables may be associated where they share a common cause. For example, smoking is a common cause (confounder) for carrying cigarette lighters and for lung cancer so there is a biasing or backdoor path from cigarette lighters to lung cancer (*cigarette lighters ← smoking → lung cancer*). Importantly, backdoor paths exist without regard to the direction of the arrows. We must control for smoking (either via statistical adjustment of study design choice) to get at the true causal effect of lighters on lung cancer.

In our example (Fig 2a), perioperative care might be used more often for older individuals, and the mortality would also be higher for older individuals so we have a red biasing path from perioperative care back through age to mortality. If we 'control' for age, indicated by shading the square grey in Fig 2a(ii), then we 'block' the backdoor path. If the graph is complete (a causal DAG), then controlling for age is sufficient to identify the true effect of perioperative care on mortality. Where important variables are unmeasured (e.g. frailty in Fig 2b), then confounding will persist.

Selection and colliders

Confounding and backdoor paths create the temptation to include all potentially related variables as controls; in addition, there is an added temptation of only selecting a population that has experienced an outcome of interest. While restricting a study population based on the presence/absence of a key confounding variable may help to control for the confounder (albeit, at the risk of decreasing generalisability), restricting a population based on the presence/absence of the outcome can introduce bias into a study. Consider investigating the relationship between smoking, asbestos exposure, and lung cancer, but only studying patients with a diagnosis of lung cancer. For a patient to have lung cancer without asbestos exposure, we would expect heavier smoking histories. Conversely, for patients to have lung cancer without smoking, we would expect greater asbestos exposure. Within this selected population, we will therefore find an inverse relationship between smoking and asbestos exposure. This subtle problem is also called Berkson's paradox.[6] In the language of graphical models, we say that lung cancer is a *collider* on the path between smoking and asbestos (*smoking → lung cancer ← asbestos exposure*) because two arrows collide onto the same variable. By knowing the participant's lung cancer status, we can detect a relationship between smoking and asbestos in any given patient, even though they are not causally connected. Therefore, conditioning the study population on the collider can introduce bias into the study.

In our example in perioperative medicine (Fig 2c), we could consider using measures of acute kidney injury (AKI) in the ICU as a variable that should be controlled for. However, AKI is likely partly genetically determined and related to standards of perioperative care. Controlling for AKI (which is a collider in the causal diagram), as shown in Fig 2c(ii), creates a biasing pathway of perioperative care and mortality through the development of AKI. Similarly, by making AKI status part of the inclusion criteria (restricting the population to only patients who experience perioperative AKI, as a way to control for AKI incidence) we would also open the same biasing path.

Causal estimation of treatment effects

These concepts allow us to inspect a graphical model and decide whether the relationship between two variables is free from systematic bias. Where we can block all backdoor paths without inducing false associations by conditioning on a collider, the relationship is said to be *identifiable*. We can simply map the study design to the graphical model to decide if we have achieved that aim. In Part 2 of this series, we will discuss examples of study designs that may help to strengthen a claim of causality between variables. Ultimately, establishing causality is both a quantitative and qualitative process. In this vein, evaluation of causality through established frameworks (such as the Bradford-Hill criteria[7]) require examination of the data through both a quantitative and qualitative lens.

Conclusion

Observational research in perioperative medicine faces threats to validity and establishing causality, primarily through two sources: spurious correlations and systematic bias. While spurious correlations can be handled through advanced statistical methods, reducing systematic bias involves knowledge of plausible biological pathways, and directionality and interplay between multiple measured and unmeasured variables. Graphical models have emerged as a tool to map out the interplay between variables, and can help strengthen the case for a causal association between an exposure and outcome variable.

Authors' contributions

Contributed to this review, focusing on at least one of the main sections: all authors.
Critical editing, additional contributions, and review of final draft: SKH, VK, KR, DM, SRM.

Declarations of interest

The authors declare that they have no conflicts of interest.

References

1. Stigler SM. *The history of statistics: the measurement of uncertainty before 1900*. Cambridge: Harvard University Press; 1986
2. Holland PW. Statistics and causal inference. *J Am Stat Assoc* 1986; **81**: 945–60
3. Pearl J. Graphical models for probabilistic and causal reasoning. In: Smets P, editor. *Quantified representation of uncertainty and imprecision*. Dordrecht: Springer; 1998. p. 367–89
4. Elwert F. Graphical causal models. In: Morgan S, editor. *Handbook of causal analysis for social research*. Dordrecht: Springer; 2013. p. 245–73

5. Khuri SF, Henderson WG, DePalma RG, Mosca C, Healey NA, Kumbhani DJ. Determinants of long-term survival after major surgery and the adverse effect of postoperative complications. *Ann Surg* 2005; **242**: 326–41
6. Moonesinghe SR, Wong DJN, Farmer L, Shawyer R, Myles PS, Harris SK. SNAP-2 EPICCS: the second sprint national anaesthesia project—EPIdemiology of critical care after surgery: protocol for an international observational cohort study. *BMJ Open* 2017; **7**. e017690–7
7. Schünemann H, Hill S, Guyatt G, Akl EA, Ahmed F. The GRADE approach and Bradford Hill's criteria for causation. *J Epidemiol Community Health* 2011; **65**: 392–5

Handling editor: Michael Avidan