BJA

British Journal of Anaesthesia, 125 (3): 398–405 (2020)

doi: 10.1016/j.bja.2020.03.032
Advance Access Publication Date: 3 May 2020
Review Article

Causal inference in perioperative medicine observational research: part 2, advanced methods

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

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

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

Summary

Although RCTs represent the gold standard in clinical research, most clinical questions cannot be answered using this technique, because of ethical considerations, time, and cost. The goal of observational research in clinical medicine is to gain insight into the relationship between a clinical exposure and patient outcome, in the absence of evidence from RCTs. Observational research offers additional benefit when compared with data from RCTs: the conclusions are often more generalisable to a heterogenous population, which may be of greater value to everyday clinical practice. In Part 2 of this methods series, we will introduce the reader to several advanced methods for supporting the case for causality between an exposure and outcome, including: mediation analysis, natural experiments, and joint effects methods.

Keywords: causal inference; confounding; epidemiology; joint effects; mediation analysis; natural experiment; observational research

Editor's key points

- The best way to test causality in science is the randomised controlled trial, but this can be expensive and infeasible.
- This article summarises three sophisticated observational statistical techniques that can be helpful in assessing causality: mediation analysis, natural experiments, and joint effects methods.
- When considering causality, it is always important to consider biological plausibility and pre-existing evidence; statistical tests in a contextual vacuum often produce misleading inferences.
- Statistical approaches, including randomised trials, cannot prove causality; they can decrease uncertainty and modify probabilistic perspectives.

The goal of observational research in clinical medicine is often to gain insight into the relationship between a clinical exposure and patient outcome, in the absence of evidence from experimental studies, such as RCTs. Clinicians often rely on observational research because of the dearth of high quality RCTs informing practice in perioperative medicine. Thus, it is paramount that observational research studies are designed as rigorously as possible, aiming to limit bias, reduce confounding, and improve patient outcomes on a global scale. With the growing availability of detailed perioperative data via the increased uptake of electronic patient records and large patient registries, a greater knowledge of advanced methods of causal inference is necessary in order to increase the rigor and reproducibility of observational findings. In this review, we build upon the concepts of graphical causal models introduced in Part 1 to introduce examples of modern observational

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

techniques (mediation analysis, natural experiments, and joint effects) and conclude with thoughts regarding claiming causality from observational studies. Although a comprehensive review of every advanced methodologic technique is well beyond the scope of this review, our goal is to give the reader an appreciation that, above and beyond the fundamentals, advanced methods are available to handle complex issues in observational research in perioperative medicine.

Mediation analysis in perioperative medicine

Generally, clinical studies are designed to assess whether an exposure (or intervention) affects a particular outcome. Although this is the most commonly instituted paradigm, sometimes either the question or the structure of data collection can answer the question: 'What is the mechanism for how an exposure affects an outcome?' For this particular question, a well-conducted mediation analysis can help to shed light on complex underlying mechanisms of the exposure–outcome relationship. In the following section, we will summarise: (1) the definition of mediating variables (and their distinction from confounding variables); (2) design consideration in a mediation study; (3) current analytic methods; (4) requirements for the consideration of mediation; and (5) limitations of this approach.

The primary conceptual framework for a mediation analysis places variables in four categories: exposures, outcomes, confounders, and mediators. Figure 1 shows an example of this concept as a directed graph, using a simplified version of a motivating example from Amato and colleagues,[1] where investigators examined whether the effect of being assigned to a 'low tidal volume' treatment group on mortality in patients with the acute respiratory distress syndrome (ARDS) was mediated by driving pressure (Fig. 1[2]). In this example, it can be seen that mediation is fundamentally different from confounding: although confounders are associated with the exposure and outcome and do not lie on the causal pathway, mediators lie on the causal pathway between exposure and outcome. From a temporal perspective, confounding variables generally occur before the exposure (i.e. particular clinical or demographic variables), and mediating variables occur after the exposure of interest. Thus, when building models to assess the relationship between an exposure and outcome, it is essential to adjust for confounders but not mediators, as this would tend to attenuate the exposure–outcome relationship. Therefore, in Figure 1,[2] the effect of low tidal volume ventilation on mortality is confounded by age, but potentially mediated through changes in driving pressure. In decomposing the associations in Figure 1,[2] the total effect of the exposure on outcome is broken up into 'direct effect' of exposure on outcome (i.e. independent of the mediator) and 'indirect effect' of exposure on outcome explained by the mediator (Fig. 2).

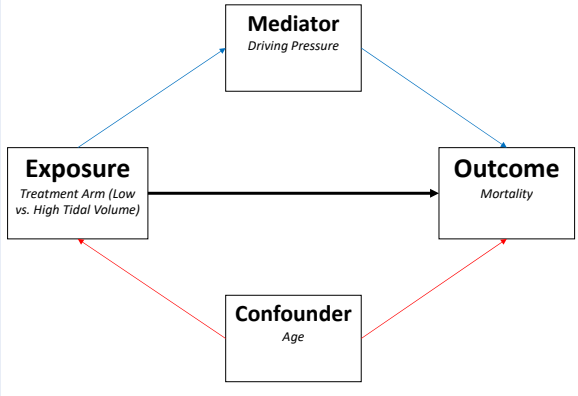

Fig 1. Conceptual model of exposure, outcome, single confounder, and single mediator.

Methodology

When designing a mediation study, several considerations need to be taken into account.[2,3] First, an appropriate conceptual model needs to be outlined, as differences between exposure, outcome, confounding, and mediating effects need to be clear – to this effect, many authors suggest that a clear causal diagram or directed acyclic graph (DAG) be constructed before analysis.[4] Second, the timing of exposure and mediating effects are a critical component, and this is generally more straightforward in acute as opposed to chronic exposures. To claim mediation, the mediator cannot occur before the exposure – this is generally more obvious in acute exposures (i.e. the effect of general anaesthesia on blood pressure) *vs* chronic exposures (i.e. the effect of smoking on myocardial infarction). Lastly, even with full optimisation, a clearly conceptualised DAG, and care in assessing the timing and biologic plausibility of the exposure–mediator–outcome relationship, cause–effect relationships with mediation analysis should be cautiously interpreted.

Traditional methods of mediation analysis involve the above conceptual framework of decomposing the exposure–outcome relationship into total effects, direct effects, and indirect effects. In epidemiologic literature, this has been termed the 'causal steps' approach to mediation analysis.[5] In the causal steps model,[6] the goal is to decompose the complex associations into a 'direct effect' (the effect of the exposure on the outcome), the association between the exposure and mediator, and the association between the mediator and outcome (controlling for the exposure), as detailed by Mascha and colleagues.[2] Using the coefficients from regression models (generally linear regression for causal steps approach, although methods are available for categorical outcomes) derived from the above equations and conceptualised in Figure 2, it is then possible to estimate the mediating (or 'indirect') effect.

To estimate the indirect effect of a mediating variable on the outcome of interest, the product of the coefficients of the exposure–mediator relationship (1) and the mediator–outcome relationship (2) can be taken; another method involves the subtraction of the direct effect (3) from the total effect (4) – generally, both methods will give the same result for continuous mediators and outcomes meeting all model assumptions.[7] If different models are used for mediator and outcome (i.e. linear regression and logit) or if different confounders are used in each model, then methods will need to be used to standardise the coefficients[8] before using the differences or product method. Although the 'indirect effect' has intrinsic meaning, for interpretation, the 'proportion mediated' is often reported, and this is given by: $1-(d/c)$. Although most statistical software packages generate standard errors for the mediator coefficients using the Sobel

Fig 2. Decomposition of total effects into direct and indirect effects.

method (using the coefficients and standard errors from the regression models), this tends to result in less precise standard errors; thus, many authors suggest an approach that uses bootstrapping,[2] which involves repeated sampling from observed data to generate a representation of the sampling distribution to derive standard errors for the coefficients in the model.

Because of the many assumptions necessary for an unbiased estimate using the causal steps approach, another approach to mediation analysis (and being used increasingly in the literature) is the potential outcomes approach, which uses a counterfactual framework to decompose the total effect into direct and indirect effects.[9–11] Using this framework, an individual can have two potential outcomes based on a binary exposure – in mediation analysis, this approach takes this principle one step further by including the mediator in the series of equations that can explain a potential outcome (the true outcome and the counterfactual situation). Thus, for an individual i, with and binary exposure E and outcome Y, we may be interested in the following effects.

Causal mediation effect (natural indirect effect)

This refers to the causal effect of the exposure that operates only through the mediator (holding exposure constant) and compares two potential outcomes: (1) mediator set to its value under the 'exposed' situation and (2) mediator set to its value in the 'unexposed' situation. Therefore, no matter what the exposure of each individual, this approach asks what the outcome would be in the presence and absence of the mediator in the counterfactual exposure situation.

Natural direct effect

Natural direct effect measures causal effect of exposure that does not operate though mediator – thus, as opposed to the above equations, in this analysis, the mediator is held constant.

Total causal effect (or total effect)

This refers to the causal effect of the exposure and mediator on outcome Y.

Because of the need to create counterfactual variables and models, a statistical software package is used to simulate potential values of the mediator and potential outcomes, and several iterations are required.

Claiming mediation requires several steps, as suggested by the evolving literature and guidance in this area.[12–14] Thus, even if a 'statistically significant' indirect effect is noted, several proposed criteria are generally required. First, it is essential that the exposure variable is associated with the mediator (i.e. effect 'a' in Fig. 2 must be significant). Second, the mediator must be associated with the outcome, independent of exposure (i.e. effect 'b' in Fig. 2 must be significant). Third, the mediation effect must be significant (both statistically and clinically), biologically plausible, and in the correct time sequence. Fourth, it is highly preferable (although sometimes the potential outcomes approach can overcome this) that no exposure–mediator interaction exists through formal testing (associations must be independent of one another). Fifth, it is best to avoid mediator–outcome confounding, in order to avoid bias in estimation of direct and indirect effects.[15] Lastly, the results of mediation analyses should be interpreted with caution, keeping in mind the limitations of mediation analyses (especially when using traditional methods) including violations of model assumptions, small sample sizes, measurement error,[16] and the use of multiple hypothesis testing. There is a growing body of literature addressing the detection of bias in mediation analyses by detecting violations to model assumption through the use of sensitivity analyses on mediated variables.[17]

Example

As discussed in the Introduction, we present a motivating example, where Amato and colleagues[1] used mediation analysis to examine the effect of changes in driving pressure (as a result of an assignment to a randomised group, based on ventilator settings) on mortality in patients with ARDS. The investigators conducted two separate analyses (one in pooled data from RCTs randomising tidal volume and another in pooled data from RCTs randomising PEEP) to examine the mediating effect of driving pressure on mortality. After examining driving pressure as a mediator candidate through stepwise tests, the investigators calculated both proportion mediated and average causal mediation effect (ACME). Among the low tidal volume trials, driving pressure mediated 75% of the benefits attributable to treatment group assignment (ACME P=0.004); and among the PEEP trials, driving pressure mediated 45% of the benefits attributable to treatment group assignment (ACME P=0.001).[1]

Mediation: summary

In summary, mediation analysis is a valuable tool for helping to decompose and characterise varying pathways from exposure to outcome; furthermore, both traditional and more modern methods can quantify these different pathways. Mediation analyses may represent unique ways to use data collected in a prospective study to help shed light on underlying disease mechanisms. Although traditional models (i.e. causal steps) are often used, limitations to those approaches have resulted in models with greater flexibility and the ability to harness the power of modern statistical software and computers to allow simulation. One must consider the many

limitations to arriving at a valid estimate from a mediation analysis, and therefore careful thought must go into deriving the appropriate conceptual model, selecting the appropriate variables and statistical model, and in carefully interpreting the results.

Natural experiment design in perioperative medicine

The history of natural experiment design can be traced to London, UK in the mid-19th century to Dr John Snow, a British physician who is credited as one of the fathers of both public health and anaesthesia. Dr Snow believed that the outbreaks of cholera that killed 616 people in London in the mid-1800s were spread by drinking water contaminated with sewage. In his 'Grand Experiment', he retrospectively compared rates of cholera infection in two populations supplied with separate water sources, one upstream from the sewage outflow into the River Thames, and the other downstream.[18] The populations were 'selected' by an influence external to any enquiry into cholera, as the majority of the scientific community believed cholera to be spread by airborne transmission or acts of God. Furthermore, it was only after the fact that Snow decided to compare these populations. Intentionally randomising people to an exposure known to be harmful would, of course, be unethical.

As is now common knowledge, cholera is indeed transmitted by water and Snow demonstrated this with a rate of 315 cholera deaths from 10 000 homes supplied with sewage-contaminated water, compared with 37 cholera deaths from 10 000 homes supplied with non-contaminated water. Although Snow's 'Grand Experiment' has since been shown to contain methodological errors, the principle of utilising external influences to 'as-if' randomise populations continues to have a great deal of utility in population health sciences.

Methodology

Natural experiments may be defined as exogenous events not controlled by the researchers, which divide a population into exposed and unexposed groups. Exposures can be environmental (natural disasters), policy (local, state, national),[19,20] medication or equipment shortages,[21] or introductions of new medications into practice. By using external influences as randomising effects, investigators may be able to reduce selection bias.[22] It remains important to ensure that study populations remain comparable, and do not differ in other unmeasured influences.

Natural experiments are useful in many situations. First, natural experiments can evaluate large scale population health interventions in an efficient and effective manner. Second, natural experiments can be exploited when randomisation is not possible; for example, an intervention deemed to be unethical. Third, given the time and cost for well-controlled RCTs (even in ethical situations), RCT evidence may not always be available to fill evidence gaps; thus, rigorously designed natural experiments may help fill this gap. Fourth, overly controlled RCTs many not mimic real-world conditions, and natural experiments may represent a method to assess the generalisability of RCT evidence in large populations.

Although the goal of a well-conducted natural experiment is to achieve 'as-if' randomisation, there are several key differences between natural experiments and traditional RCTs.[23] Although both RCTs and natural experiments have clearly defined interventions, the details of the intervention in an RCT are within full control of the investigators, including dose, timing, compliance, etc. Furthermore, assignment of the intervention in RCTs is also under control of the researchers, so that appropriate randomisation techniques can be used. In contrast, natural experiments do not allow detailed knowledge of the assignment process, thereby subjecting the analysis to confounding. Lastly, all subjects in an RCT have a known risk of receiving a treatment or control assignment; but in natural experiment designs, the possibility of a subject being 'at risk' for the exposure and outcome remain unclear. Based on this, natural experiments cannot be seen as a substitute for a well-designed RCT, but rather as complementary in the quest for efficacy, generalisability, and causality.

There are several approaches to evaluating natural experiments, above and beyond descriptive assessment before and after intervention.[22] The use of traditional regression methods can allow for inclusion of measured confounding variables in an analysis, using time period (before and after intervention) as the exposure. In order for natural experiment study participants to be allocated to either an intervention or control group while minimising measured and unmeasured confounding, propensity scoring can be additionally used. The propensity score is defined as the probability of a study subject being allocated to the intervention group, and is most commonly calculated using logistic regression. The derived propensity score can be used for matching, inverse probability of treatment weighting, or as covariate in a regression model. The intention of this method is that by accounting for the measurable differences in subjects assigned to the intervention and control groups, unmeasured confounding will also be minimised. However, the limitations of propensity score matching are related to the number and type of available covariates, and how those covariates are related to unmeasured confounders.[23] Furthermore, although RCTs theoretically assure balance of measured and unmeasured confounders, propensity scores only assure relative balance of measured confounders. In addition to examining only a population that has been exposed to an intervention (before and after intervention), including a control group that has been unexposed to an intervention using difference-in-differences methodology[22] may help to further enhance causal inference. Lastly, the inclusion of detailed time series data is generally more robust than only measuring data at a single time point, so that changes in trends (i.e. slopes) can be analysed before and after intervention — this methodologic approach has been described as an interrupted time series[24] and can be evaluated with advanced analytic methods, such as segmented regression.[25]

Example

We present a motivating example, in which Vail and colleagues[21] exploited a putative natural experiment to examine the association between a norepinephrine shortage in the USA and mortality in patients with septic shock.

In the primary analysis, the investigators examined 27 835 patients treated in 26 'shortage' hospitals (where norepinephrine utilisation for vasopressor support in septic shock was observed to decrease during the period of the US norepinephrine drug shortage). Among patients treated for septic shock at 'shortage' hospitals, in-hospital mortality was observed to be 35.9% during non-shortage time periods, compared with 39.6% during the period of the US

norepinephrine drug shortage (adjusted odds ratio=1.15; 95% confidence interval [CI], 1.01–1.30; P=0.03). To account for secular trends and further support causal inference, the investigators also constructed a difference-in-differences model (using 'consistent use' hospitals, where norepinephrine utilisation was not observed to change during the drug shortage period, as a control group), with stable results (adjusted odds ratio=1.17; 95% CI, 1.06–1.31; P=0.003). The investigators concluded that patients with septic shock admitted to 'shortage' hospitals during the period of the US norepinephrine shortage had higher in-hospital mortality, compared with admission during non-shortage periods.

Natural experiments: summary

Natural experiment design can offer a viable alternative to RCTs, where RCTs are impractical because of logistical or ethical considerations. 'As-if' randomising study subjects by an external influence has the potential to minimise unmeasured confounding, particularly if paired with additional strategies such as regression adjustment, propensity score matching, use of control groups, and statistical methods for time series data. However, despite the use of rigorous methodologic and analytic techniques, the observational nature of natural experiment designs may still be prone to the threat of unmeasured confounding.

Joint effects in perioperative medicine

Sometimes the effect of one drug on an outcome differs with the presence of another drug to some extent, resulting in a heterogeneity of treatment effects. For example, when examining the effect of multimodal analgesia on outcomes after surgery, certain analgesic combinations may be synergistic, whereas others may not (or worse, antagonistic). In this situation, drawing a causal association between an exposure (Drug A) and outcome may depend on the presence/absence of an additional exposure (Drug B). In other words, an interaction may exist between two drugs (or exposures), and outcome effects are thus not independent. In epidemiology papers or textbooks, this is called an interaction or 'joint effect' between the two exposures. If the joint effect of two exposures is larger or smaller than the sum of the individual effects, there is interaction. In many cases a joint effect is assessed by entering a product term of two exposures into the regression model. Two types of outcomes, binary and continuous outcomes, are often used in clinical studies, and the interpretation of the regression coefficient of the product term varies based on which outcome is used. In general, in joint effect analysis, linear regression models using a continuous outcome estimate additive interaction, whereas logistic regression models (and Cox proportional hazard models) using a binary outcome estimate multiplicative interaction. For the sake of simplicity, we will focus only on interaction on additive scales using a continuous outcome, and we will discuss the methods and formulas to estimate joint effect with a hypothetical example.

Methodology

As a motivating example, consider hypothetical data concerning the effect of two multimodal analgesics on opioid consumption on the day of surgery. It is of interest to determine whether the joint effect of two multimodal analgesics would have lower opioid consumption after surgery than the sum of individual effects. Let Y denote a continuous outcome: opioid consumption. Let A and B denote two binary exposures of interest. These can be acetaminophen, NSAIDs, gabapentinoids, etc. We can construct a 2×2 table with the mean of opioid consumption in the four groups: non-A and non-B (A−B−), A and non-B (A+B−), non-A and B (A−B+), and A and B (A+B+) in Table 1.

In practice, joint effect is often evaluated by including a product term for the two exposures in the model. A linear regression model in observational studies adjusting for covariates might take the form:

$$Y=\beta_0+\beta_1 A+\beta_2 B+\beta_3 AB$$

where AB simply denotes the product of the two exposures, and β_0, β_1, β_2, and β_3 are the corresponding regression coefficients. The mean values of the outcome for the four groups defined by A and B in Table 1 will be as follows. The individual effect of A is: $Y_{A+B-}=\beta_0+\beta_1$. The individual effect of B is: $Y_{A-B+}=\beta_0+\beta_2$. The combined effect (doubly exposed) is: $Y_{A+B+}=\beta_0+\beta_1+\beta_2+\beta_3$. The effect of no A and no B (doubly unexposed) is: $Y_{A-B-}=\beta_0$. The assessment of interactions using the interaction contrast (IC) is to measure the extent to which the combined effect of the two exposures exceeds the sum of the individual effects.[26] In other words, the IC is a measure of the difference between the observed Y_{A+B+} estimate and the expected Y_{A+B+} estimate (based on individual effects).[27] This could be measured by:

$$\begin{aligned}\text{IC} &= \text{observed}(Y_{A+B+}) - \text{expected}(Y_{A+B+}) \\ &= (Y_{A+B+} - Y_{A-B-}) - [(Y_{A+B-} - Y_{A-B-}) + (Y_{A-B+} - Y_{A-B-})] \\ &= (\beta_0+\beta_1+\beta_2+\beta_3-\beta_0) - [(\beta_0+\beta_1-\beta_0)+(\beta_0+\beta_2-\beta_0)] \\ &= \beta_3\end{aligned}$$

If the IC ($=\beta_3$) is not zero, we conclude that there is interaction. Positive interaction is defined by the extent to which the combined effect is larger than the sum of their individual effects, whereas negative interaction is defined by the extent to which the combined effect is less than the sum of their individual effects (Fig. 3).[28–30]

In addition, in joint effect analysis clinicians would be more interested in whether there is synergism or antagonism when two drugs are administered, rather than just to say there is positive or negative interaction. However, although the terms synergism and antagonism have specific meanings, they are often used ambiguously.[28,30] It is generally straightforward to evaluate synergism or antagonism when both exposures affect outcomes in the same direction. However, for opposing individual effects ($\beta_1>0$ and $\beta_2<0$, or $\beta_1<0$ and $\beta_2>0$), there are cases in which the definition of synergism (or antagonism) varies based on the magnitude of the effects of exposures.[31] For simplicity, if we assume that two exposures affect the outcome in the same direction,[28] we can define synergism and

Table 1 Interaction model for joint effect of A and B.

	B−	B+
A−	β_0	$\beta_0+\beta_2$
A+	$\beta_0+\beta_1$	$\beta_0+\beta_1+\beta_2+\beta_3$

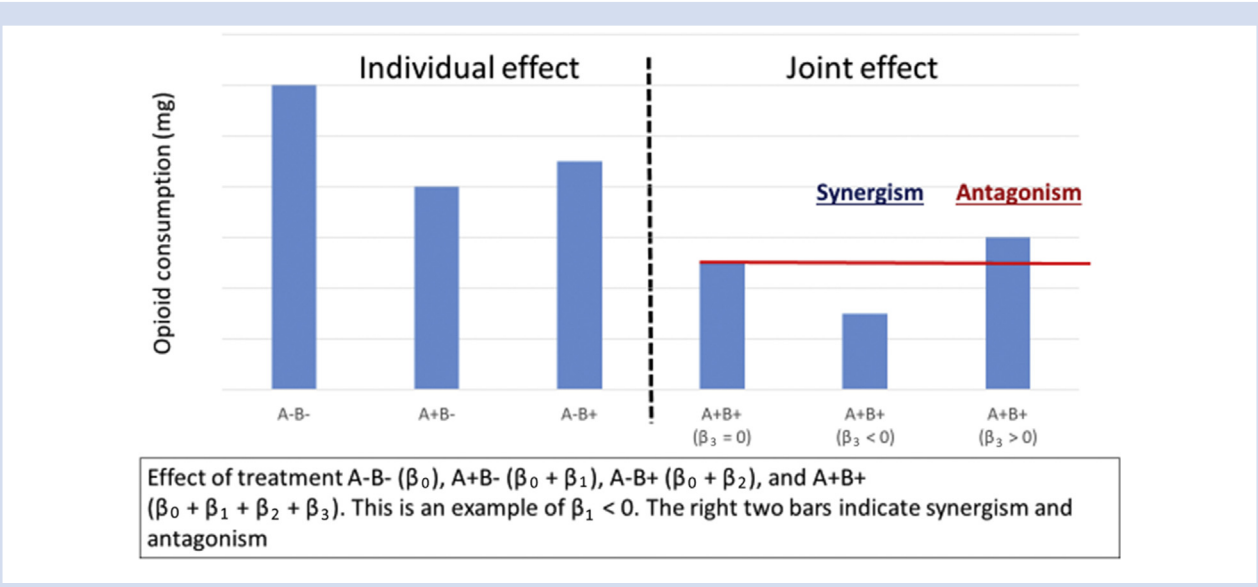

Fig 3. A diagram for conceptualising the joint effect of drug A and B on opioid consumption. Synergism is when the effect of [A+B+] > A+B, and antagonism is when the effect of [A+B+] < A+B.

antagonism as follows. When a primary exposure is positively associated with the outcome ($\beta_1>0$), IC<0 indicates antagonism and IC>0 indicates synergism. In contrast, when a primary exposure is inversely associated with the outcome ($\beta_1<0$), IC<0 indicates synergism and IC>0 indicates antagonism. If IC=0 (the 95% CI of β_3 includes 0), the combined effect of A and B is defined as having insufficient evidence of interaction.[27]

Example

Multimodal analgesia plays a critical role with the goal of reducing opioid use and associated side-effects. Multimodal analgesics commonly include regional analgesia, NSAIDs, selective cyclooxynegase-2 inhibitors, acetaminophen, gabapentinoids, and ketamine, which can all lead to reduction of acute pain through differing receptor pathways. A key question is whether specific combinations of multimodal analgesics have joint effects when they are used together.[32] In a hypothetical dataset, consider NSAIDs and acetaminophen as 'risk factors' for decreased opioid consumption on the day of

Table 2 Example of synergism. Coefficients of a linear regression model with NSAIDs and acetaminophen as binary exposures and product of NSAIDs and acetaminophen in a hypothetical dataset. Outcome is opioid consumption on the day of surgery (mg). CI, confidence interval.

Parameter	Coefficient	Estimate	Lower CI	Upper CI
NSAIDs	β_1	−2.0	−2.1	−1.9
Acetaminophen	β_2	−5.0	−5.1	−4.9
NSAIDs × Acetaminophen	β_3	−3.0	−4.6	−1.4
Constant	β_0	70	69.0	71.0

Table 3 Example of antagonism. Coefficients of a linear regression model with NSAIDs and acetaminophen as binary exposures and product of NSAIDs and acetaminophen in a hypothetical dataset. Outcome is opioid consumption on the day of surgery. CI, confidence interval.

Parameter	Coefficient	Estimate	Lower CI	Upper CI
NSAIDs	β_1	−2.0	−2.1	−1.9
Acetaminophen	β_2	−5.0	−5.1	−4.9
NSAIDs × acetaminophen	β_3	3.0	1.4	4.6
Constant	β_0	70	69.0	71.0

surgery. NSAIDs and acetaminophen are binary variables, and opioid consumption is a continuous variable. NSAIDs, acetaminophen, and a product term of NSAIDs and acetaminophen are entered in a linear regression model. Examples of synergism and antagonism are presented in Tables 2 and 3. In practice, β_3 is directly obtained by statistical software so that the IC is readily available in this setting. Table 2 shows negative interaction because of IC<0 and synergism because of β_1<0 and IC<0. Table 2 shows positive interaction because of IC>0 and antagonism because of β_1<0 and IC>0.

Joint effects: summary

We have discussed interaction analyses in terms of the measures, estimation procedures, and interpretation. The methods and formulas are intended to assist clinical researchers who plan to evaluate joint effect between two binary exposures on a continuous outcome. These tools are based on classical theories of interaction analysis. We have not been able to describe all methods related to interaction analysis, but we hope that the reader will consult the relevant papers and we also hope that this section provided useful information on how to perform and interpret analyses of joint effects.

Conclusion

To appropriately utilise the vast amount of data available to clinical researchers, and to fill gaps in knowledge where high-quality evidence from RCTs are lacking, rigorous observational research designs continue to have an important role in clinical medicine and public health. We have discussed several methods to improve causal inference using observational research designs in perioperative medicine. With growing data, statistical knowledge, and computing power, we expect further techniques to emerge in the coming decades. Despite the proliferation of methodologic and analytic techniques to help support causality, the clinician should remember that causal inference is both a quantitative and qualitative exercise. With a solid understanding of the fundamentals, and utilisation of advanced techniques in clinical research, observational research in perioperative medicine will continue to advance at a rapid pace.

Authors' contributions

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

Declarations of interest

The authors declare that they have no conflicts of interest.

References

1. Amato MB, Meade MO, Slutsky AS, et al. Driving pressure and survival in the acute respiratory distress syndrome. *N Engl J Med* 2015; **372**: 747–55
2. Mascha EJ, Dalton JE, Kurz A, Saager L. Statistical grand rounds: understanding the mechanism: mediation analysis in randomized and nonrandomized studies. *Anesth Analg* 2013; **117**: 980–94
3. Turan A, Mascha EJ, Roberman D, et al. Smoking and perioperative outcomes. *Anesthesiology* 2011; **114**: 837–46
4. Akinkugbe AA, Sharma S, Ohrbach R, Slade GD, Poole C. Directed acyclic graphs for oral disease research. *J Dent Res* 2016; **95**: 853–9
5. Krause MR, Serlin RC, Ward SE, Rony RY, Ezenwa MO, Naab F. Testing mediation in nursing research: beyond Baron and Kenny. *Nurs Res* 2010; **59**: 288–94
6. Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 1986; **51**: 1173–82
7. Alwin DF, Hauser RM. The decomposition of effects in path analysis. *Am Sociol Rev* 1975; **40**: 37–47
8. Hayes AF. Beyond Baron and Kenny: statistical mediation analysis in the new millennium. *Commun Monogr* 2009; **76**: 408–20
9. Maldonado G, Greenland S. Estimating causal effects. *Int J Epidemiol* 2002; **31**: 422–9
10. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. *Psychol Methods* 2010; **15**: 309–34
11. Imai K, Keele L, Yamamoto T. Identification, inference, and sensitivity analysis for causal mediation effects. *Stat Sci* 2010; **25**: 51–71
12. Shrout PEBN. Mediation in experimental and nonexperimental studies. *Psychol Methods* 2002; **7**: 422–45
13. Albert JM. Mediation analysis via potential outcomes models. *Stat Med* 2008; **27**: 1282–304
14. Collins LGJ, Flaherty B. An alternative framework for defining mediation. *Multivariate Behav Res* 1998; **33**: 295–312
15. Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: methods, interpretation and bias. *Int J Epidemiol* 2013; **42**: 1511–9
16. le Cessie S, Debeij J, Rosendaal FR, Cannegieter SC, Vandenbroucke JP. Quantification of bias in direct effects estimates due to different types of measurement error in the mediator. *Epidemiology* 2012; **23**: 551–60
17. VanderWeele TJ. Bias formulas for sensitivity analysis for direct and indirect effects. *Epidemiology* 2010; **21**: 540–51
18. Snow J. On the mode of communication of cholera. *Edinb Med J* 1856; **1**: 668–70
19. Haghpanahan H, Lewsey J, Mackay DF, et al. An evaluation of the effects of lowering blood alcohol concentration limits for drivers on the rates of road traffic accidents and alcohol consumption: a natural experiment. *Lancet* 2019; **393**: 321–9
20. Vandoros S, Avendano M, Kawachi I. The EU referendum and mental health in the short term: a natural experiment using antidepressant prescriptions in England. *J Epidemiol Community Health* 2019; **73**: 168–75
21. Vail E, Gershengorn HB, Hua M, Walkey AJ, Rubenfeld G, Wunsch H. Association between US norepinephrine shortage and mortality among patients with septic shock. *JAMA* 2017; **317**: 1433–42
22. Craig P, Katikireddi SV, Leyland A, Popham F. Natural experiments: an overview of methods, approaches, and contributions to public health intervention research. *Annu Rev Public Health* 2017; **38**: 39–56
23. Streiner DL, Norman GR. The pros and cons of propensity scores. *Chest* 2012; **142**: 1380–2

24. Leatherdale ST. Natural experiment methodology for research: a review of how different methods can support real-world research. *Int J Soc Res* 2018; **22**: 19–35
25. Mascha EJ, Sessler DI. Segmented regression and difference-in-difference methods: assessing the impact of systemic changes in health care. *Anesth Analg* 2019; **129**: 618–33
26. Greenland S, Lash TL, Rothman KJ. Concepts of interaction. In: Rothman KJ, Greenland S, Lash TL, editors. *Modern epidemiology*. Philadelphia: Wolters Kluwer Health; 2008. p. 71–86
27. Knol MJ, van der Tweel I, Grobbee DE, Numans ME, Geerlings MI. Estimating interaction on an additive scale between continuous determinants in a logistic regression model. *Int J Epidemiol* 2007; **36**: 1111–8
28. VanderWeele TJ, Knol MJ. Remarks on antagonism. *Am J Epidemiol* 2011; **173**: 1140–7
29. Choi B, Ostergren PO, Canivet C, et al. Synergistic interaction effect between job control and social support at work on general psychological distress. *Int Arch Occup Environ Health* 2011; **84**: 77–89
30. VanderWeele TJ, Robins JM. The identification of synergism in the sufficient-component-cause framework. *Epidemiology* 2007; **18**: 329–39
31. Piggott JJ, Townsend CR, Matthaei CD. Reconceptualizing synergism and antagonism among multiple stressors. *Ecol Evol* 2015; **5**: 1538–47
32. Feldheiser A, Aziz O, Baldini G, et al. Enhanced Recovery after Surgery (ERAS) for gastrointestinal surgery: Part 2. Consensus statement for anaesthesia practice. *Acta Anaesthesiol Scand* 2016; **60**: 289–334

Handling editor: Michael Avidan