



HHS Public Access

Author manuscript

J Surg Res. Author manuscript; available in PMC 2023 October 01.

Published in final edited form as:

J Surg Res. 2022 October ; 278: 155–160. doi:10.1016/j.jss.2022.04.059.

A clinician's guide to conducting research on causal effects

Vivian H. Lyons, PhD, MPH^{a,b}, Jamaica RM Robinson, PhD, MPH^c, Brianna Mills, MA, PhD^{b,d}, Elizabeth Y Killien, MD, MPH^{b,e}, Stephen J. Mooney, PhD^{b,d}

^aDepartment of Health Behavior and Health Education, University of Michigan, Ann Arbor, MI, USA;

^bHarborview Injury Prevention & Research Center, University of Washington, Seattle, WA, USA;

^cDepartment of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA;

^dDepartment of Epidemiology, School of Public Health, University of Washington, Seattle, WA, USA;

^eDivision of Pediatric Critical Care Medicine, Department of Pediatrics, University of Washington, Seattle, WA, USA

Abstract

Surgeons are uniquely poised to conduct research to improve patient care, yet a gap often exists between the clinician's desire to guide patient care with causal evidence and having adequate training necessary to produce causal evidence. This guide aims to address this gap by providing clinically relevant examples to illustrate necessary assumptions required for clinical research to produce causal estimates.

Keywords

Causality; Epidemiological Research Design; Clinical Research; Academic Surgery

Introduction

Surgeons working at the intersection of academics and clinical care are uniquely poised to conduct informed, relevant, and timely research to improve processes of care and patient outcomes. Research involvement can also broaden surgeons' perspectives on their clinical work, satisfy intellectual curiosity, and aid in career development.^{1,2} Additionally, research is often an explicit expectation of many clinical training programs and an essential component of academic promotion criteria.

Address correspondence to: Elizabeth Y. Killien, MD MPH, Seattle Children's Hospital, Pediatric Critical Care Medicine FA 2.112, 4800 Sand Point Way NE, Seattle, WA 98105, elizabeth.killien@seattlechildrens.org, (206) 987-5838.

Author contributions: Drs. Lyons, Robinson, Mills, and Mooney conceptualized the project, provided background intellectual content, and drafted the initial manuscript. Dr. Killien provided clinical context and reviewed and revised the manuscript. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Conflicts of Interest: The authors have no conflicts of interest to declare.

While most medical students and academic physicians believe that participation in clinical research is important,^{3–6} very few have formal training in research methodology⁵ or sufficient knowledge of biostatistics to conduct rigorous research.⁷ The number of clinicians participating in formal physician-scientist pathways or intensive Master's or doctoral degree programs is declining.⁸ Most clinicians first conduct research during non-degree-awarding academic residencies, fellowships, and other positions that may not include formal research training.

It is important to ensure that clinicians who have not undertaken additional advanced research training have the resources and support they need to conduct high quality clinical research. As epidemiologists working in an interdisciplinary injury research center, we often find ourselves working with surgeons who are caught between a desire to implement the best possible analytic strategy for a given research project and uncertainty around which strategy is the most appropriate. One of the most frequent challenges we have observed is confusion regarding causality in clinical research. This confusion includes understanding whether a research *question* is causal, when a statistical estimate may be interpreted as a *causal effect*, and how to best contribute to the evidence base when available data and statistical estimates aren't suited to interpretation as causal effects.

Clinicians without formal research training may rely on simplifying rules-of-thumb regarding evidential value in clinical research. One such rule is that only randomized controlled trials produce causal evidence.^{9,10} Yet, in practice, clinical decisions are informed by observations comprising the best available evidence rather than only evidence from randomized trials.^{11,12} For example, one of the most common daily decisions surgeons must make is selection of appropriate analgesia for their patients while inpatient and upon discharge. While opioids have long been a mainstay of perioperative pain management, many hospitals have increased their scrutiny of opioid prescribing given concern about the worsening opioid epidemic. As a provider who must balance your patients' analgesia requirements with these public health concerns, you would like to know whether opioid prescriptions after acute trauma or surgery are contributing to opioid misuse. You know that this isn't a question that can be answered in a randomized controlled trial, but can you still assess it for a causal link? Our objective is for this guide to serve as an entry point to causal inference for a surgeon or other clinical practitioner with basic statistical knowledge and a research question about a potentially causal relationship.

What is a causal research question?

In general, epidemiologists roughly categorize clinical research questions as (1) descriptive, (2) predictive, or (3) causal. Understanding this taxonomy and how your research question fits into it can help you to select an appropriate analytic approach and linguistic framing for your project (Table 1).

1. Descriptive research characterizes distributions of disease prevalence, risk factors, or outcomes in a specific population, often within a specific time window. Findings from descriptive research provide a foundation for generating and refining hypotheses for future research endeavors, while also informing policymaking.¹³ For example, you might design a study comparing the mean

number of opioid prescriptions filled after hospitalizations for traumatic injury at different trauma facilities with rates of opioid abuse in that region. Descriptive studies may include comparisons (such as prevalence of opioid prescriptions at discharge by hospital type) but they do not support a counterfactual outcome (e.g., they do not ask how risk factor or disease distribution would differ if characteristics of the population or interventions were different).

2. Predictive research questions employ clinical data to predict outcomes for an individual patient or patient population given what is already known about that patient or population. For example, a clinical decision support algorithm developed from predictive research might forecast that a 65-year-old with a femur fracture with ongoing unresolved pain at discharge who is prescribed 30 days of opioids has a 20% one-year risk of developing long-term prescription opioid use. Unlike descriptive questions, which examine the present or the past, predictive research forecasts a specific future for an individual or defined population. However, like descriptive research, predictive research does not try to determine how the disease course or condition would change as the result of a different treatment choice.
3. Causal research questions ask how changes in health status result from changes in exposure or treatment. For example, a surgeon who asks if they should prescribe only non-steroidal anti-inflammatory drugs (NSAIDs) at discharge rather than opioids to discourage long-term opioid use is asking a causal question – if they change their behavior, will it cause a change in outcome? Or specifically, does an opioid prescription after traumatic injury contribute to risk of opioid use disorder? A key characteristic of this type of question is its counterfactual contrast.¹⁴ Even though as a clinical researcher you observe, at most, one outcome for each individual (i.e., the patient developed long-term prescription opioid use or they did not develop long-term prescription opioid use), you are interested in projecting what the outcome would have been had the exposure been different than what it was; that is, if it were *counter to fact* (i.e., if the individual had taken non-opioid pain management versus the short course of prescribed opioids).

Ultimately, all three types of research are important and provide evidence for clinical decisions. However, the third research type, causal research, is the only type that demonstrates a direct effect of an intervention and is frequently the most challenging to conduct and interpret.

What is required for research to be causal?

Statistical analyses from any population-based study, including both observational studies and randomized trials, will typically estimate a controlled association between a treatment and an outcome. For example, to assess whether a history of prescribed opioids following an injury is associated with a higher risk of opioid overdose, your analysis might use statistics to hold every other measured patient characteristic (e.g., gender, age, baseline health status) constant, and identify that those with a history of prescribed opioids had five times the rate of opioid overdose when compared with those without a history of prescribed opioids.

Is this five-fold elevated risk the effect of prescribed opioids on opioid overdose incidence in the population that your study data comes from? Not necessarily. Even with a causal research question and a perfectly conducted research study of any design, a statistical parameter is not guaranteed to accurately estimate the population average causal effect. The plausibility that a statistical parameter (e.g. the five-fold risk observed above) represents a causal effect depends on a set of core assumptions.¹⁵

Core assumptions—There are three core causal inference assumptions: (1) consistency, (2) positivity, and (3) conditional exchangeability.

1. **Consistency** is the assumption that your exposure, treatment, or intervention of interest is applied equally to all individuals classified as exposed, and not applied at all to individuals classified as unexposed.¹⁶ If you wanted to compare outcomes among patients prescribed opioids to those prescribed NSAIDs at discharge, you might be concerned the consistency assumption would be violated due to variation in the specific opioid prescribed, the daily dose, the time interval between doses, the duration of the prescription, and how the prescribed dose changed over time.
2. **Positivity** is the assumption that there could be both exposed and unexposed people in each group of covariates on which you analytically stratify (e.g. age, gender, medical history), such that you are able to describe the distribution of the outcome across exposure levels in each covariate group.¹⁷ For example, suppose your study was evaluating opioid use disorder incidence among patients initially prescribed opioids, adjusted for hospital and insurance status. If one of the included hospitals had a policy to prescribe lower cost NSAIDs rather than opioids to patients lacking health insurance, uninsured patients in the hospital that never prescribed opioids to uninsured patients would be systematically precluded from exposure status, violating the positivity assumption.
3. **Conditional exchangeability** is the assumption that, before treatment, exposed and unexposed individuals have equivalent probability of the outcome (conditional on covariates that have been controlled for).¹⁴ In a study assessing whether opioid prescription use leads to increased risk of opioid dependency, it would help to have demographic information (e.g., age at prescription opioid initiation) and/or health status data (e.g., chronic pain) from participants. Satisfying conditional exchangeability requires an in-depth understanding of prior literature and theoretical frameworks that describe how relevant covariates influence the relationship between your exposure and outcome of interest.

If you have a causal research question, do you need to conduct a randomized controlled trial (RCT)?

In short: it's nice if you can, but it's not necessary. Clinicians and health researchers typically consider RCTs to be at the top of the 'evidence pyramid,' with good reason. Randomized controlled trials are designed to generate data where exposure (treatment) allocation meets the consistency assumption – by specifying the intervention that individuals receive, investigators hope to minimize differences in exposure to the point of being

ignorable. Assigning the intervention usually also allows an RCT to meet the positivity assumption (every participant has a chance of being allocated the exposure) and, when randomization succeeds, the conditional exchangeability assumption (on average, the exposed group has the same predilection for the outcome as the unexposed group, except for the impact of the exposure). In short, RCTs are designed to increase the probability that those core assumptions will hold, allowing interpretation of statistical parameters as causal effects.

However, conducting an RCT may be unfeasible for an array of reasons, including lacking necessary financial or time resources, or having an exposure or hypothesis that is not possible or ethical to apply and/or alter for trial participants. It is also possible that the sample of willing participants may not be sufficiently representative of the broader patient population to produce meaningful results.¹⁸ For example, you may consider enrolling patients into an RCT for an experimental opioid tapering protocol, but are concerned that patients prone to opioid use disorder would systematically decline to participate in the trial, which would result in estimating an effect that would not translate to the actual population of interest. Importantly, even RCTs are not guaranteed to meet core assumptions required for causal research.¹⁹

If you have a causal research question and do not conduct an RCT, what makes a statistical parameter estimate interpretable as a causal effect?

If your study cannot meet the core assumptions of consistency, positivity, and conditional exchangeability – and observational studies usually cannot – your statistical estimates cannot be interpreted as causal effects. However, even if your estimates are not causal *effects*, they still can provide causal *evidence*. Evidence from controlled associations obtained with descriptive research provides foundational evidence for future causal hypotheses and research, and also may lead to clinical changes that can themselves be assessed more rigorously.^{10,12} Consider that most of the evidence that smoking causes lung cancer is associational – as detailed previously, there are almost always issues with violation of all three core causal inference assumptions – but there are no plausible alternate causes for that relationship other than a causal effect of smoking.

Furthermore, consider that many of the methods used to approach a causal research question are neither necessary nor sufficient for answering the research question by themselves but do often provide valuable context for better understanding of the research question.²⁰ A Directed Acyclic Graph (DAG) is one such tool, which is used to graphically visualize the hypothesized causal relationships between the exposure, the outcome, and all related covariates. For example, say you want to estimate the impact of instituting a tapering protocol on opioid prescriptions and subsequent opioid dependency. Suppose you know that at your center, younger age is associated with being included in the tapering protocol, and that age may also affect the risk of developing an opioid addiction. Using this information to draw a DAG would not only illustrate the relationship hypothesized by the research question between the exposure and outcome of interest, but also the “back door path” through patient’s age that connects the opioid tapering protocol to opioid dependency (Figure 1).²¹ Since the focus of a DAG is on covariates that influence both the exposure and outcome,

visualizing the research question via this method will also help you be parsimonious with the number of covariates to consider and include in statistical analyses.²⁰ A good place to start learning about this visualization method is [DAGitty.net](https://dagitty.net).²²

Additionally, even if your associational estimates cannot be interpreted as causal effects, you may be able to perform additional sensitivity or quantitative bias analyses to bolster the causal evidence. For example, suppose you conduct a randomized trial of an experimental opioid taper protocol, but you are only able to follow up 90% of your participants to the trial's end point. Because the actual effect of the opioid taper depends on the outcomes of the full 100% of participants, you cannot directly interpret your statistical parameter as the estimated causal effect of the taper. However, you could conduct secondary analyses that explore what the statistical parameter would be if everyone who was lost to follow up developed long-term opioid use and what the parameter would be if nobody who was lost to follow up developed long-term opioid use. These analyses would thus place bounds on the impact of your loss to follow-up. This is one example of the broader field of *quantitative bias analysis*, which is an analytic approach to exploring how much error would need to be present in a study to meaningfully change the appropriate interpretation of findings.^{23,24}

More broadly, if you are careful in how you refer to the association you estimated, your discussion *can* interpret your results in light of your causal question of interest. For example, in a multisite study evaluating the association between opioid prescriptions after abdominal trauma and ongoing pain at two-week follow-up, you might be concerned that referral patterns affect the severity of abdominal injuries treated at hospitals with different solid organ injury management protocols beyond what can be accounted for statistically using injury severity scores. You should then report the association you observe, but also remind readers whether or not your association is consistent with your hypothesis that opioid prescriptions are not associated with pain two weeks after injury. Many analyses can support this approach, including instrumental variables, inverse probability weighting, or targeted maximum likelihood estimation, among others.^{25–27}

What do you do if you are not confident your analysis can produce an estimate of a causal effect?

When you have a causal research question, it is appropriate to use causal language throughout your writing to describe your *question* and underlying hypothesis. You would *like* to know whether your exposure causes your outcome. However, when you cannot interpret your estimates as causal effects, you should ensure that the language you use to report your findings does not imply that your study produced such an estimate.²⁸

For instance, the word “effect” is used to denote the causal impact of an exposure on an outcome; if your statistical parameter cannot be interpreted as a causal effect, you can still describe what you actually estimated, which was “the association between” your exposure and outcome. Table 2 contains some easy substitutions for causal language, which can be applied to your results and discussion sections.

In short, it is important to be precise about both the question you would like to know the answer to (e.g. will prescribing NSAIDs rather than opioids achieve adequate pain control?)

and the evidence you actually constructed (e.g. people who received NSAIDs reported adequate pain control and fewer side effects than people who received opioids, even after statistical control for injury type and age).

Conclusion

Causality is at the heart of clinical decision-making, yet formal causal evidence is frequently unavailable to contribute to these decisions. A clinical researcher filling gaps in the evidence typically seeks an answer to a causal question. In practice, that clinician might be unable to conduct an RCT due to resource, ethical, or logistic barriers. Yet any clinical evidence can be useful when it comprises the best available answer to the question, with precision, accuracy, and acknowledgment of limitations. An understanding of the causal assumptions can help identify and articulate these limitations. When possible, partnering *early* in the research process with collaborators trained in study design can help develop appropriate research designs, and ensure planned research activities are designed to allow estimation of the desired parameter.

Acknowledgments:

The authors would like to thank the Harborview Injury Prevention & Research Center as well as Dr. Anjum Hajat, Dr. Ali Rowhani-Rahbar, and Dr. Marco Carone for their input and support in developing this manuscript.

Funding:

This work did not receive specific funding but authors were supported by the following grants during manuscript preparation: National Center for Advancing Translational Sciences of the National Institutes of Health (TL1 TR002318), National Institute of Child Health and Human Development (T32HD057822, K23HD100566), National Library of Medicine (K99LM012868), National Cancer Institute (T32CA094880, T32CA094061), National Institute of Environmental Health Sciences (T32ES015459), and the Firearm Safety Among Children & Teens Consortium funded by the National Institute for Child Health and Human Development (1R24HD087149).

Role of Funder/Sponsor:

The NIH had no role in the design, writing, or submission of the work. The content is solely the responsibility of the authors and does not represent the official views of the National Institutes of Health.

References

1. Jain MK, Cheung VG, Utz PJ, et al. Saving the Endangered Physician-Scientist - A Plan for Accelerating Medical Breakthroughs. *N Engl J Med*. 2019;381(5):399–402. [PubMed: 31365796]
2. Rahman S, Majumder MA, Shaban SF, et al. Physician participation in clinical research and trials: issues and approaches. *Advances in medical education and practice*. 2011;2:85–93. [PubMed: 23745079]
3. Paget SP, Caldwell PH, Murphy J, et al. Moving beyond ‘not enough time’: factors influencing paediatric clinicians’ participation in research. *Internal medicine journal*. 2017;47(3):299–306. [PubMed: 27925382]
4. Stone C, Dogbey GY, Klenzak S, et al. Contemporary global perspectives of medical students on research during undergraduate medical education: a systematic literature review. *Medical education online*. 2018;23(1):1537430.
5. Sumi E, Murayama T, Yokode M. A survey of attitudes toward clinical research among physicians at Kyoto University Hospital. *BMC medical education*. 2009;9:75. [PubMed: 20025782]
6. Yeh HC, Bertram A, Brancati FL, et al. Perceptions of division directors in general internal medicine about the importance of and support for scholarly work done by clinician-educators.

- Academic medicine : journal of the Association of American Medical Colleges. 2015;90(2):203–208. [PubMed: 25374036]
7. West CP, Ficalora RD. Clinician attitudes toward biostatistics. *Mayo Clinic proceedings*. 2007;82(8):939–943. [PubMed: 17673062]
 8. Kosik RO, Tran DT, Fan AP, et al. Physician Scientist Training in the United States: A Survey of the Current Literature. *Evaluation & the health professions*. 2016;39(1):3–20. [PubMed: 24686746]
 9. Scriven M A Summative Evaluation of RCT Methodology: & An Alternative Approach to Causal Research. *Journal of MultiDisciplinary Evaluation*. 2008;5(9):11–24.
 10. Vandembroucke JP, Broadbent A, Pearce N. Causality and causal inference in epidemiology: the need for a pluralistic approach. *International journal of epidemiology*. 2016;45(6):1776–1786. [PubMed: 26800751]
 11. Cartwright N Are RCTs the Gold Standard? *BioSocieties*. 2007;2(1):11–20.
 12. Sackett DL. Evidence-based medicine. *Seminars in perinatology*. 1997;21(1):3–5. [PubMed: 9190027]
 13. Kaufman JS. There is no virtue in vagueness: Comment on: Causal Identification: A Charge of Epidemiology in Danger of Marginalization by Sharon Schwartz, Nicolle M. Gatto, and Ulka B. Campbell. *Annals of epidemiology*. 2016;26(10):683–684. [PubMed: 27641315]
 14. Hernán MA, Robins JM. Estimating causal effects from epidemiological data. *Journal of epidemiology and community health*. 2006;60(7):578–586. [PubMed: 16790829]
 15. Petersen ML, van der Laan MJ. Causal models and learning from data: integrating causal modeling and statistical estimation. *Epidemiology (Cambridge, Mass)*. 2014;25(3):418–426.
 16. Cole SR, Frangakis CE. The consistency statement in causal inference: a definition or an assumption? *Epidemiology (Cambridge, Mass)*. 2009;20(1):3–5.
 17. Westreich D, Cole SR. Invited commentary: positivity in practice. *American journal of epidemiology*. 2010;171(6):674–677; discussion 678–681. [PubMed: 20139125]
 18. Westreich D, Edwards JK, Lesko CR, et al. Target Validity and the Hierarchy of Study Designs. *American journal of epidemiology*. 2019;188(2):438–443. [PubMed: 30299451]
 19. Deaton A, Cartwright N. Understanding and misunderstanding randomized controlled trials. *Soc Sci Med*. 2018;210:2–21. [PubMed: 29331519]
 20. Pearce N, Lawlor DA. Causal inference—so much more than statistics. *International journal of epidemiology*. 2016;45(6):1895–1903. [PubMed: 28204514]
 21. NIDA. What is the scope of prescription drug misuse in the United States? National Institute on Drug Abuse website. January 26, 2022. <https://nida.nih.gov/publications/research-reports/misuse-prescription-drugs/what-scope-prescription-drug-misuse>. Accessed March 2, 2022.
 22. Textor J, van der Zander B, Gilthorpe MS, et al. Robust causal inference using directed acyclic graphs: the R package ‘dagitty’. *International journal of epidemiology*. 2016;45(6):1887–1894. [PubMed: 28089956]
 23. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Annals of internal medicine*. 2017;167(4):268–274. [PubMed: 28693043]
 24. Lash TL FM, Fink AK. *Applying quantitative bias analysis to epidemiologic data* Vol 192. New York: Springer; 2009.
 25. Hogan JW, Lancaster T. Instrumental variables and inverse probability weighting for causal inference from longitudinal observational studies. *Statistical methods in medical research*. 2004;13(1):17–48. [PubMed: 14746439]
 26. Ohlsson H, Kendler KS. Applying Causal Inference Methods in Psychiatric Epidemiology: A Review. *JAMA psychiatry*. 2020;77(6):637–644. [PubMed: 31825494]
 27. Schuler MS, Rose S. Targeted Maximum Likelihood Estimation for Causal Inference in Observational Studies. *American journal of epidemiology*. 2017;185(1):65–73. [PubMed: 27941068]
 28. Petitti DB. Associations are not effects. *American journal of epidemiology*. 1991;133(2):101–102. [PubMed: 1985440]

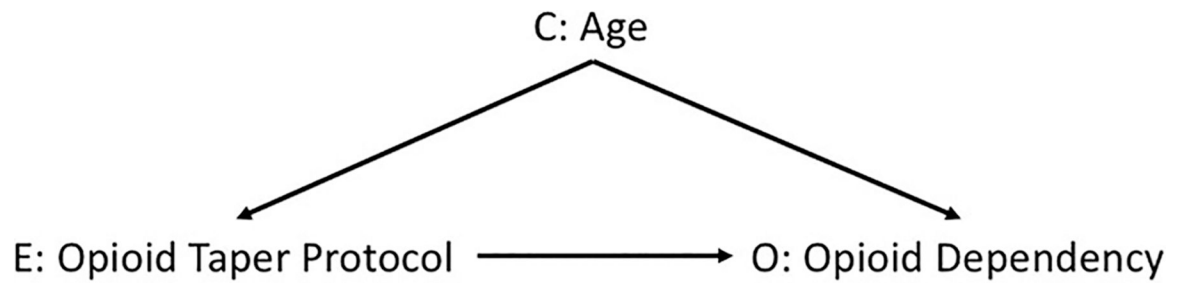


Figure 1: Example of a Directed Acyclic Graph (DAG) of an opioid taper protocol and dependency.

E indicates the exposure variable, O indicates the outcome of interest, and C indicates a confounding variable.

Table 1.

Categories of clinical research

Question type	Defining characteristic	Example
Descriptive	Examine the world as it is	What percentage of opioid dependency started with a post-injury prescription?
Predictive	Describe a future world	Based on the last five years of post-injury prescription rates, will opioid dependency rates go up or down in the next five years?
Causal	Estimate the impact of some change in the world	Would substituting NSAIDs for opioids among bone fracture patients reduce opioid use disorder?

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Language substitutions for causal/non-causal research

Causal Language	Non-Causal Language
caused	was associated with
increased/decreased	was associated with an increase/decrease, had higher/had lower
the effect of X on Y	the association between X and Y

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript