

# A Narrative Review of Methods for Causal Inference and Associated Educational Resources

Douglas Landsittel, PhD; Avantika Srivastava, MS; Kristin Kropf, MS

**Background and Objectives:** Root cause analysis involves evaluation of causal relationships between exposures (or interventions) and adverse outcomes, such as identification of direct (eg, medication orders missed) and root causes (eg, clinician's fatigue and workload) of adverse rare events. To assess causality requires either randomization or sophisticated methods applied to carefully designed observational studies. In most cases, randomized trials are not feasible in the context of root cause analysis. Using observational data for causal inference, however, presents many challenges in both the design and analysis stages. Methods for observational causal inference often fall outside the toolbox of even well-trained statisticians, thus necessitating workforce training. **Methods:** This article synthesizes the key concepts and statistical perspectives for causal inference, and describes available educational resources, with a focus on observational clinical data. The target audience for this review is clinical researchers with training in fundamental statistics or epidemiology, and statisticians collaborating with those researchers. **Results:** The available literature includes a number of textbooks and thousands of review articles. However, using this literature for independent study or clinical training programs is extremely challenging for numerous reasons. First, the published articles often assume an advanced technical background with different notations and terminology. Second, they may be written from any number of perspectives across statistics, epidemiology, computer science, or philosophy. Third, the methods are rapidly expanding and thus difficult to capture within traditional publications. Fourth, even the most fundamental aspects of causal inference (eg, framing the causal question as a target trial) often receive little or no coverage. This review presents an overview of (1) key concepts and frameworks for causal inference and (2) online documents that are publicly available for better assisting researchers to gain the necessary perspectives for functioning effectively within a multidisciplinary team. **Conclusion:** A familiarity with causal inference methods can help risk managers empirically verify, from observed events, the true causes of adverse sentinel events.

**Key words:** causality, confounding, potential outcomes, propensity scores, secondary data

**Author Affiliation:** Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, Pennsylvania.

**Correspondence:** Douglas Landsittel, PhD, Department of Biomedical Informatics, University of Pittsburgh, The Offices at Baum, 5th Floor, 5607 Baum Blvd, Pittsburgh, PA 15206 (dpl112@pitt.edu).

The authors thank all collaborators who contributed to the development of the motivating educational resources, including (1) the "Decision Tool for Causal Inference and Observational Data Analysis Methods in Comparative Effectiveness Research," (2) "An Online Self-Guided Course for Writing a Successful Concept Proposal in Patient-Centered Outcomes Research," and (3) "An Online Self-Guided Course in Propensity-Score Based Methods for Causal Inference"; the full list of acknowledgments for each project can be found under the corresponding link for each project at the authors' Web site ([www.landsittelab.pitt.edu/educational-resources](http://www.landsittelab.pitt.edu/educational-resources)). All statements in this article, including its findings and conclusions, are solely those of the authors and do not necessarily represent the views of the Agency for Healthcare Research and Quality, the National Libraries of Medicine, or the Patient-Centered Outcomes Research Institute, its Board of Governors or Methodology Committee.

The development of the above-referenced educational resources was supported through funding by the Patient-Centered Outcomes Research Institute (contract #R-IMC-1306-03827; PI: Landsittel), the Agency for Healthcare Research and Quality (grant #R25HS023185; PI: Landsittel), and an administrative supplement to the National Libraries of Medicine-funded University of Pittsburgh Biomedical Informatics Training Program (grant #5 T15 LM007059-32; PI: Hochheiser). Landsittel was funded by all 3 funding sources; Srivastava and Kropf were funded only by the AHRQ grant #R25HS023185.

The authors declare no conflicts of interest.

Q Manage Health Care  
Vol. 29, No. 4, pp. 260–269

Copyright © 2020 Wolters Kluwer Health, Inc. All rights reserved.

DOI: 10.1097/QMH.0000000000000276

Root cause analysis (RCA) can be defined as the process for identifying causal factor(s) for an incident or variation in process.<sup>1-3</sup> While RCA is often conceptualized as an investigative process, RCA may also be accomplished through analyzing observational data, for example, assessing causal effects for adverse outcomes from electronic health records (EHRs). The increasing availability of EHR data will continue to facilitate the use of large observational data sets for RCA.<sup>4</sup>

Analysis of EHRs for RCA requires the establishment of causal relationships between exposures (or interventions) and outcomes. Establishing causality necessitates either randomization or sophisticated methods applied to well-designed observational studies.<sup>5,6</sup> Otherwise, observed statistical relationships will likely reflect association, not causation. Since randomized trials are, in general, neither applicable nor feasible for RCA, investigators must carefully plan observational studies and use methods that optimally estimate unbiased causal effects specific to a given research question. Doing so represents a challenging problem, as numerous confounding factors typically exist between potential causes and outcomes.

The primary challenge for establishing causal rather than associative relationships with observational data is the presence of confounding (ie, factors related to both the exposure and the outcome). While multivariable methods, such as regression analysis, are often used to adjust for potential confounders, such methods do not

directly estimate causal effects. Furthermore, to establish causality, observational studies must be rigorously designed in a way that captures temporal relationships and potential confounders.<sup>7,8</sup> This review summarizes the scope of methods and designs that can estimate causation versus association and highlights strategies for associated workforce training.

## METHODS

The objective of this narrative review is to outline and describe the scope and complexity of methods for causal inference, with a focus on clinical applications amiable to RCA using observational studies of secondary data sources. The “Results” section is organized as follows.

1. Definitions and frameworks for establishing causality.
2. Key aspects of observational study designs required as a prerequisite for potentially establishing causality with secondary observational data.
3. The processes of using a target trial and causal road map to emulate a randomized trial for estimating causal effects.
4. The concept and variations of propensity score-based methods, which are one possible approach for the statistical model (as one step of the causal road map).
5. A review of other statistical models used for estimating the causal effect.
6. A description of available educational resources, including textbooks, review articles and tutorials, and other educational resources.

The “Discussion” section then summarizes the complexity of necessary skills and the subsequent implications for workforce training.

## RESULTS

### Definitions and frameworks for causality

This section outlines several common approaches to define a framework for causal inference.

#### Hill's criteria

Hill's criteria are often referenced as the standard definition for causality in epidemiology (with almost 10 000 citations).<sup>9</sup> Hill proposes 9 criteria, using the example of assessing causality between smoking exposure and lung cancer.<sup>10</sup>

1. *Strength*: Is the exposure associated with a large effect in worsening the outcome?
2. *Consistency*: Have study findings been repeated across different circumstances?
3. *Specificity*: Can the exposure be linked to a specific illness(es) or cause(s) of mortality?
4. *Temporality*: Is the exposure measured before the incidence of the outcome?
5. *Biological gradient*: Is there a dose-response relationship between exposure and illness severity?
6. *Plausibility*: Is a causal relationship biologically plausible?

7. *Coherence*: Is the hypothesized causal relationship consistent with existing knowledge?
8. *Experiment*: Can the causal relationship be verified through experimentation?
9. *Analogy*: Can we define an analogy between the hypothesized effect and another accepted causal relationship?

Hill's criteria represent a set of considerations but do not provide a specific mathematical definition of causality. Without such a definition for causal effects, the question of whether a given statistical approach estimates an association versus causation remains difficult to evaluate, thus motivating other frameworks for causality.

#### The potential outcomes framework

The potential outcomes framework<sup>11</sup> has become an increasingly popular way to define causality. The general idea, which can be traced back to the work of Neyman,<sup>12,13</sup> is to define 2 hypothetical outcomes of a given experiment, where, for a given subject, one outcome is observed and the other is unobserved (and referred to as the counterfactual outcome). If, for instance, subject  $i$  receives either medication A or medication B, the potential outcomes are denoted by  $Y_i^{(A)}$  and  $Y_i^{(B)}$ . All circumstances surrounding the subject are exactly the same except the medication received. The individual causal effect, or equivalently the individual treatment effect, is then defined as a statistical contrast (eg, difference or ratio or other comparisons) between  $Y_i^{(A)}$  and  $Y_i^{(B)}$ . However, since only 1 observation is measured, the individual causal effect cannot be directly estimated.

Under randomization, we can consistently estimate the expected contrast, where any difference in subject characteristics is purely due to chance. However, this does not hold for observational data, where the treatments or exposures received typically depend on the subjects' characteristics or other factors. Differences in characteristics between the exposed and unexposed subjects, therefore, cannot be attributed to chance, and the statistical contrast between outcomes from the 2 groups is subsequently a biased estimate of the causal effect.

For observational data, the potential outcomes framework represents a useful definition for distinguishing between association and causation. Specifically, the estimation of treatment effects with traditional regression models yields the conditional expectation of  $Y$ , given a set of explanatory variables (including treatment status). This conditional expectation is mathematically different from the marginal expectation defined as the causal effect based on potential outcomes. Other methods, such as propensity score-based methods, are therefore needed to estimate causal effects, although significant debate exists on the feasibility of necessary assumptions.<sup>14</sup>

#### Other frameworks for establishing causality

Although this review focuses on the potential outcomes framework to define causality, it is worth noting

that a number of other frameworks exist. Two examples are directed acyclic graphs<sup>15</sup> (and related Bayesian methods<sup>16</sup>) and the sufficient component cause framework.<sup>17,18</sup> Directed acyclic graphs, which graphically display the hypothesized or observed relationships between variables, are becoming increasingly common in medical research<sup>19,20</sup> and causal discovery<sup>21</sup> (ie, searching for causal relationships consistent with empirical data). Several publications describe the relationships between these frameworks.<sup>22,23</sup>

### Study design and assumptions for establishing causality with secondary data

The need for carefully designed and transparently reported randomized trials is well known and supported through guidance documents and required reporting.<sup>24,25</sup> These steps serve to optimize reproducibility of results and produce well-designed studies that achieve consistent (eg, asymptotical unbiased) estimates of causality. While similar efforts for planning and reporting of observational studies exist,<sup>26</sup> a more prevailing opinion for secondary data (which has already been collected for a different purpose) can be “we can only analyze what we have,” thus creating the misconception that study design is not relevant for secondary data.<sup>8</sup>

In fact, observational studies, even for secondary data, involve far more complexities in the design phase. As emphasized by Rubin,<sup>7</sup> “for objective causal inference, design trumps analysis.” More specifically, the complexities in design of secondary data analyses can be divided into 2 main issues of (1) evaluating data sufficiency for addressing causal questions, and (2) further constraining the data to meet causal inference assumptions. The second issue is addressed in the next section as part of defining a target trial.

Before using secondary data for causal inference, the sufficiency of the data must be examined relative to the causal question. Questions of data sufficiency include the following:

1. Are data collected in a manner that capture temporal associations, with confounding factors measured prior to exposures and exposures measured prior to outcome?
2. Are sufficient data available to quantify confounding factors?
3. Do all subjects have some nonzero probability of being either exposed or unexposed?
4. Do all subjects have the same set of possible outcomes across the possible exposure levels, and are the potential outcomes of different subjects independent of each other?

Each of these requirements represents a necessary prerequisite to estimate causal effects in a consistent manner. Temporality is an explicit component of causality (as described by Hill) since confounders and exposures occurring at the same time as outcomes likely reflect associations. Measuring confounders is necessary for achieving conditional independence of the potential outcome and the exposure (also referred to as ignorability)<sup>27</sup> based on some statistical model. The

third requirement, which is referred to as positivity,<sup>28</sup> is necessary to avoid confounded by indication. The final requirement corresponds to the stable unit-treatment value assumption, which is a general assumption for causal inference.<sup>29</sup>

### Defining causal effects using a target trial and causal road map

This section describes the approach of using a target trial and causal road map for making causal inferences based on observational data; these concepts are often neglected in analyzing observational data but are critical for valid causal inference.

#### The target trial

Hernán and Robins<sup>30</sup> propose framing observational studies as a hypothetical randomized trial using inclusion and exclusion criteria specific to the clinical question. Data should also have well-defined treatment strategies (or exposure measurement) and be able to specify the process behind participant follow-up and outcome measurement.

This approach leads to asking well-defined and estimable causal questions, such as “does assigning physicians to working longer shifts lead to increased medication errors and increased mortality?” versus poorly defined questions such as “is it better to have shorter shifts for physicians?” This approach also serves to meet the above-described assumptions. For instance, if some subset of the study participants always receive the intervention, they would be excluded from the target trial. The target trial should also be defined so that the measured data are sufficient to account for potential confounding and temporal associations.

Another component of the target trial is defining the causal effect. For RCA, the causal effect of interest is typically the per-protocol or average treatment effect among the treated (ATT). Other variations (which are less applicable to RCA) include the intention-to-treat or overall average treatment effect (ATE), or the complier average causal effect. Different designs and different causal inference methods correspond to different causal estimands<sup>31</sup> (ie, different statistical contrasts of the potential outcomes being estimated). The causal effect therefore needs to be specified, implemented, and interpreted in a consistent manner with the research question; this issue is often ignored in the literature when applying causal inference methods.

#### The causal road map

Peterson and van der Laan<sup>32,33</sup> propose using the following 7-step process:

1. Describe knowledge through a causal model and/or causal graph, for example, directed acyclic graphs and associated structural equations,<sup>34</sup> for quantifying the nature of variables involved in the causal relationships.
2. Specify the observed data and their relationship to the causal model.

3. Define the target causal quantity in terms of an explicit function of the potential outcomes.
4. Use the causal model and observed data to assess identifiability, that is, whether the question can be expressed statistically. If not, describe needed assumptions or additional data.
5. Specify a statistical model and associated estimand to evaluate the clinical question.
6. Conduct the estimation procedure.
7. Use all of the above steps, and related assumptions, to interpret results.

The above steps, along with the specification of the problem as, and restriction of the data set to, a target trial, provide a set of strategies for using observational data to evaluate causal effects about an exposure. The following 2 sections describe relevant statistical models.

**Propensity score–based methods**

Propensity score–based methods represent one set of approaches for specifying the statistical model used to estimate the causal effect (ie, for steps 5-6 in the causal road map). These methods are usually specific to point exposures (ie, received at a single time point), although some variations may be used for time-varying exposures.<sup>35</sup> This review refers to these models as “propensity score–based” (PS) methods to emphasize the variability in associated approaches. Similar to the use of regression models in traditional statistics, PS methods are not a single method but rather a set of methods and associated strategies.

The PS methods can be motivated by first considering a randomized trial, where factors that predict differences in exposure are, on average, identically distributed between exposure groups. For observational studies, however, numerous factors can significantly affect each subject’s propensity for being exposed. The PS methods seek to statistically model that propensity and then create a new sample that pseudo-randomizes the data so that both exposure groups have similarly distributed propensities. Pseudo-randomization is

usually accomplished via matching on the estimated propensity (and discarding unmatched subjects), stratifying on the propensities (and assuming homogeneity within strata), or weighting the data on the propensity (similar to reweighting probability samples to represent the original population). Once the data are pseudo-randomized, they can be analyzed via standard statistical methods. These analyses are captured through the following steps.

1. Specify the causal estimand, for example, ATE or ATT. The choice of the causal effect depends on the clinical question and needs to be considered in step 3.
2. Specify a model for the assignment mechanism to estimate each subject’s propensity for being exposed; examples include a standard logistic or a more complex machine learning model. The choice for the optimal model is difficult to determine and depends on the complexity of the underlying relationship and relative trust in the data versus the hypothesized model and its assumptions.
3. Pseudo-randomize the sample based on matching,<sup>36-39</sup> stratification,<sup>40</sup> or inverse probability of treatment weighting.<sup>41</sup> There are numerous variations of each approach<sup>42,43</sup>; the approach, and associated algorithm, must be consistent with the causal effect of interest.
4. Estimate the final treatment effect by applying an outcomes model (usually a traditional unadjusted statistical test or regression model) that is consistent with both the outcome distribution and the pseudo-randomization approach.

The Figure visually displays and elaborates on the above steps with a description of the associated challenges and a summary of the associated variations in the modeling process.

Steps 1, 3, and 4 are interconnected, as different pseudo-randomization approaches apply to different estimands, and require a different outcomes model. For instance, matching estimates the ATT, whereas PS

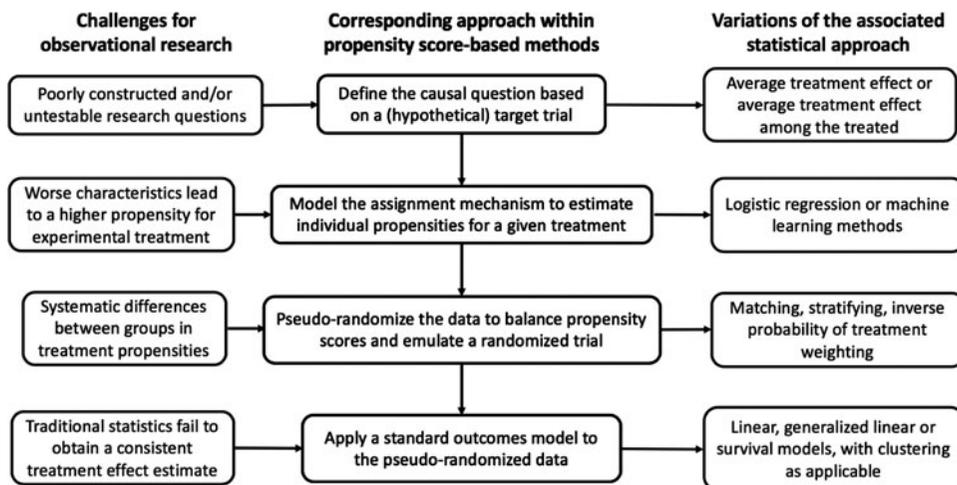


Figure. Components and variations of propensity score–based methods.

weighting usually estimates the ATE (although variations of weighting can estimate the ATT). The outcomes model also needs to account for the matching (eg, paired *t* test or conditional logistic regression), stratifying, or weighting (eg, treating the PS estimates as survey weights with the same methods as used for probability samples).

Since publication of the seminal papers on PS methods,<sup>39,40</sup> tens of thousands of publications have used or described their use. The PS methods represent probably the most popular choice for the statistical model and estimation of causal effects; a search on “propensity scores” and “causal inference” (in Google Scholar between 1980 and 2019) yields more than 11 000 results. In the first half of 2020 alone, a large number of tutorials and review articles have been published across a range of disciplines, including journals specific to rheumatology,<sup>44–46</sup> clinical epidemiology,<sup>47</sup> cardiovascular medicine,<sup>48</sup> trauma surgery,<sup>49</sup> emergency medicine,<sup>50</sup> pharmaceutical research,<sup>51–56</sup> neurosurgery,<sup>57</sup> transplantation,<sup>58,59</sup> athletic training,<sup>60</sup> oncology,<sup>61–63</sup> obstetrics,<sup>64,65</sup> anesthesia,<sup>66</sup> aging,<sup>67</sup> addiction,<sup>68,69</sup> pediatrics,<sup>70</sup> and more general medical applications.<sup>71–77</sup> A number of textbooks<sup>78–82</sup> have also been written specific to PS methods. Despite the substantial volume of published literature, the complexity of such methods (eg, as characterized by the many options for each step in the Figure) is still poorly understood and often incorrectly applied in the literature.<sup>83–89</sup>

### Other statistical models for causal inference

The PS methods represent only one possible approach for the statistical model; describing the details of PS methods serves to highlight the complexities and assumptions that are part of any statistical model. The choice of statistical model depends on many considerations, including the following:

1. Is the exposure or intervention received at a specific time point, or is it a time-varying exposure (or intervention sustained over time)?
2. Is the estimand consistent with the causal question of interest?
3. How well do a model’s key assumptions agree with the observed data?
4. Should we use a nonparametric model that makes less assumptions about the nature of relationships, or use a parametric model that may use the data more efficiently?
5. Can we achieve conditional independence between the exposure and outcome by conditioning on observed covariates, or do we need to employ alternative strategies that account for unmeasured confounding?

Choice of the estimand is particularly critical for RCA, since estimates of average causal effects are less relevant for root causes. The questions about assumptions for a given statistical model are difficult to answer and often not directly testable. Answering the above 5 questions will not specifically answer “which method

works best” but rather provides guidance as to which methods should be considered or favored.

### Variations of propensity score methods

Other variations of PS methods include doubly robust and covariate balancing PS (CBPS) methods. Doubly robust approaches separately model the relationships between confounders and the outcome within each exposure group to separately predict each of the potential outcomes.<sup>90</sup> Doubly robust methods may improve statistical properties of the subsequent causal estimates<sup>91</sup> by requiring only correct specification of *either* the outcome model *or* the PS model for the exposure mechanism. The CBPS method<sup>92</sup> simultaneously maximizes (1) the conditional probability of exposure, given measured covariates, and (2) the balance of covariates between exposure groups. The CBPS methods are also more robust to deviations from the specified PS model.

In terms of the above-specified questions (ie, exposure timing, choice of estimand, feasibility of assumptions, including conditional independence, and use parametric vs nonparametric models), doubly robust and CBPS methods generally apply to point exposures but may offer significant advantages in terms of the assumptions related to specification of the PS model. Attention, however, still needs to be given to the choice of the estimand, as it can differ depending on the selected variation. Both doubly robust and CBPS models may use variations with either parametric or nonparametric methods.

### Network models and structural equations

Creating a causal graph (ie, set of directed relationships between exposures and outcomes) can help identify appropriate variables and associated relationships for the statistical models. Those relationships may include confounding, effect modification, or mediation of the exposure–outcome relationships. Causal diagrams thus represent an important step in specifying the statistical model for almost any causal inference approach, but these approaches can also be used for specifying structural equations and directly estimating causal effects.<sup>93</sup>

Pearl<sup>34</sup> provides an extensive review, including connections between structural equation models, causal graphs, and the potential outcomes model. In particular, Pearl describes approaches to estimate causality by blocking paths between the exposure and the outcome. While such methods are still not as widely used in clinical research, more than 5000 articles have been published (from searching “causal graph”) since the initial landmark publications in 1986. In terms of the above-specified questions (on timing, estimand, assumptions, and parametric vs nonparametric models), causal graphs and structural equation methods can apply to both point and time-varying exposures. The assumptions needed depend on the variations of the approach being used, which can be either parametric or nonparametric.

### **Instrumental variables**

Instrumental variables (IV) are variables that predict exposure but are otherwise independent of the outcome<sup>94,95</sup>; randomization is the ultimate instrument, since randomization should predict who receives the intervention, but is otherwise independent of the outcome, including being independent of any other measured or unmeasured factors that are potentially associated with the outcome. The IV approach then seeks to use one of several methods<sup>96,97</sup> to quantify variation in the IV as a type of natural experiment for estimating causal effects. As an example, Hearst et al<sup>98,99</sup> use a subject's draft status for the Vietnam War as the IV to assess the causal effects of serving in the war on postwar mortality.

In terms of the above-specified questions, IV methods also generally apply to point exposures. Instrumental variables methods may offer advantages over other methods since they do not require measuring the confounding variables but instead essentially create a natural experiment to estimate causal effects. However, other assumptions are critical for valid inference, including the strength of the instrument and its conditional independence from the outcome. The estimand, which is usually the complier average causal effect, also differs from PS-based methods. Instrumental variables methods may use variations with either parametric or nonparametric methods.

### **Other complex methods for time-varying exposures and sustained interventions**

While most applications in the literature focus on point exposures or interventions, RCA may depend on time-varying exposures (eg, specific to conditions for treating physician) or sustained interventions (eg, medications that vary over time and depend on the outcome). Dissecting causal effects from these scenarios can be far more complex than the case of point exposures or interventions.<sup>100</sup> For instance, antiretroviral treatment of HIV<sup>101</sup> to increase CD4 count can vary over time depending on viral load, but viral load also directly affects CD4 count. Either ignoring or adjusting for CD4 count therefore leads to either confounding (without adjustment) or overadjusting as the effect of treatment is mediated by CD4 count. Standard approaches and PS methods will therefore be invalid for this type of problem.

There are, however, several approaches<sup>102</sup> available for evaluating causality in this scenario, including g-methods,<sup>103-106</sup> which essentially standardize the data, marginal structural models,<sup>35,107</sup> which use PS weights applied over multiple time points, modifications of IV approaches,<sup>108</sup> and targeted maximum likelihood estimation methods.<sup>109,110</sup> In terms of the above-specified questions, these methods extend more basic causal inference approaches to time-varying exposures but either have stronger parametric assumptions or require more extensive data to model complex longitudinal relationships in a nonparametric fashion (eg, with the g-computation formula<sup>111</sup>). The estimand also differs across these methods.

### **Educational resources**

A number of relevant textbooks have been written, including textbooks on PS methods,<sup>78-82</sup> matching approaches,<sup>112</sup> mediation and interactions,<sup>113</sup> targeted maximum likelihood estimation and machine learning,<sup>114,115</sup> and causal graphs and/or associated structural equation models.<sup>116-122</sup> Other textbooks provide general coverage across a range of causal inference methods,<sup>123-131</sup> several of which are specific to the social sciences.<sup>132-135</sup> Others provide some coverage of causal inference, but in the larger context of epidemiology and health sciences,<sup>136-138</sup> or comparative effectiveness research<sup>139</sup> and quasi-experimental designs,<sup>140</sup> or big data,<sup>141</sup> temporal data,<sup>142</sup> and/or regression modeling.<sup>143</sup> Thousands of tutorials have been published; a search on tutorial and "causal inference" yields more than 3000 publications in the last 4 years alone.

Several researchers and organizations have also created Web sites, videos, and/or other resources for causal inference. Online training includes a causal inference seminar series,<sup>144</sup> at least 4 different Coursera courses specifically on causal inference<sup>145</sup> (attended by nearly 40 000 students), and other causal inference-related resources,<sup>146-153</sup> including links to software packages and extensive references. A few of the Web-based resources include guidance that can be useful for designing and conducting causal inference analyses with secondary data.

1. The Patient-Centered Outcomes Research Institute (PCORI) has a Methods Program that funds proposals to improve patient-centered outcomes research (PCOR), which includes causal inference. Their portfolio can be found online<sup>154</sup> and includes (currently 116) funded studies focused on matching methods, subgroup analyses, study designs, comparisons with more than 2 treatments, machine-learning methods, new IV methods, and variable selection for PS analysis of rare outcomes. Completed projects have a link to a final report and publications.
2. PCORI also has a Research Methodology Web page<sup>155</sup> with links to Methodology Standards. Several of the standards are directly relevant to causal inference, including topics of data integrity and rigorous analyses, data registries and networks, and a category specifically on causal inference. The Web page also includes a link to their academic curriculum, which has a series of video tutorials for each of the topics.
3. "Comparative Effectiveness Research Based on Observational Data to Emulate a Target Trial"<sup>153</sup> (CERBOT) is a Web-based program with an introduction and 5 different modules to assist users through the process of specifying a target trial.
4. The "Decision Tool for Causal Inference and Observational Methods and Data Analysis Methods in Comparative Effectiveness Research" (DECODE CER)<sup>156</sup> is an online tool in Google Drive that provides a set of links for each step of the causal inference process, from specifying the

research question, to assessing data adequacy and assumptions, modeling the assignment mechanism, pseudo-randomizing the sample, fitting the outcomes model, and specifying IV and estimating effects. DECODE CER also includes a data extraction of 168 articles from a systematic review of statistical properties of causal inference methods, a link to other causal inference resources, and other general information on comparative effectiveness.

5. "An Online Self-Guided Course in Propensity Score-Based Methods for Causal Inference"<sup>157</sup> is another online tool in Google Drive that provides a course for writing a project proposal and analysis plan using PS methods for observational causal inference. The initial landing page links to a syllabus-like document with a description of the course and links to 8 modules. These modules include (1) general concepts, including potential outcomes, stating the research question, and assessing study designs, (2) steps of the PS method, and (3) assessing the impact of unmeasured confounding.
6. The "Center for Causal Discovery" Web site<sup>158</sup> describes applications of causal discovery and includes a number of training videos and software packages and programs for creating causal graphs and using causal discovery algorithms.

In summary, a large volume of published literature and other educational resources are available to researchers interested in using secondary observational data for causal inference and RCA. While many of the articles and textbooks are highly technical, other tutorials and Web-based video resources are appropriate for a more general audience, with a stress on the underlying concepts necessary to facilitate collaboration within a multidisciplinary team.

## DISCUSSION

The increasing availability of data from large EHR systems, across one or multiple institutions, will continue to facilitate use of secondary data for RCA. While doing so can lead to improved discovery of factors that lead to adverse events, use of observational secondary data also carries significant risk for making incorrect conclusions about causal relationships. Valid use of appropriate causal inference methods is therefore critical to RCA with secondary data.

The complexity of using causal inference methods requires that statisticians and/or epidemiologists collaborate effectively with clinical researchers in a multidisciplinary team to develop a target trial and causal question and select and implement methods consistent with the research objectives along each step of the causal road map. Selection of the appropriate statistical model must also consider which assumptions can be reasonably supported by the available data, either directly or after introducing restrictions as part of the target trial design. The causal estimand associated with a given method also needs to be consistent with

the research question. Furthermore, when comparing results across multiple studies, the choice of methods and the resulting causal estimand need to be considered; otherwise, apparent differences may be the result of comparing apples with oranges.

Further enhancing workforce training is critical to improving the quality of research and validity of findings for causal relationships. Although use of causal inference methods is becoming increasingly common among statisticians and epidemiologists, most graduate programs still lack formal training in these concepts and methods. Instead, most practitioners gain expertise strictly through independent study, later professional training, and/or on-the-job experience. In addition to statisticians and epidemiologists, clinical researchers are also increasingly interested in using and receiving training in these methods. Doing so, however, represents a significant challenge because of the complexities of not just the statistical models (which is only one step of the causal road map) but also the complexity of the concepts, the number of different types of approaches available for causality, and the differences in assumptions and causal estimands.

## CONCLUSIONS

Causal inference methods are complex and require knowledge of (1) the overriding conceptual frameworks (eg, specifying the target trial and steps in the causal road map) and scope of methods, (2) an understanding of relevant study design issues, and (3) skills in the statistical methods necessary for selecting and conducting analyses with the optimal statistical models (which are only one step of the causal road map). To address these challenges and take full advantage of causal inference methods for RCA using secondary data, researchers should form multidisciplinary teams with both clinical and statistical expertise. This review provides a summary of the relevant concepts and scope of methods; any associated gaps in knowledge across the research team can be addressed by (1) using the CERBOT and/or DECODE CER tools to gain a clear understanding of the general process, (2) consulting the training videos on the Center for Causal Discovery Web site for formulating causal graphs, (3) using the online course on PS methods to develop an analysis plan, if PS methods are applicable, (4) further reviewing the PCORI resources and/or other literature for planning the design and/or other statistical models, and (5) specifying a target trial and causal road map to make valid causal inferences.

## REFERENCES

1. Wu AW, Lipshutz AK, Pronovost PJ. Effectiveness and efficiency of root cause analysis in medicine. *JAMA*. 2008;299(6):685-687.
2. Peerally MF, Carr S, Waring J, Dixon-Woods M. The problem with root cause analysis. *BMJ Qual Saf*. 2017;26(5):417-422.
3. Neily J, Ogrinc G, Mills P, et al. Using aggregate root cause analysis to improve patient safety. *Jt Comm J Qual Saf*. 2003; 29(8):434-439.

4. Davis Giardina T, King BJ, Ignaczak AP, et al. Root cause analysis reports help identify common factors in delayed diagnosis and treatment of outpatients. *Health Aff.* 2013;32(8):1368-1375.
5. Greenland S. Randomization, statistics, and causal inference. *Epidemiology.* 1990 1(6):421-429.
6. Rubin DB. Causal inference using potential outcomes: design, modeling, decisions. *J Am Stat Assoc.* 2005;100(469):322-331.
7. Rubin DB. For objective causal inference, design trumps analysis. *Ann Appl Stat.* 2008;2(3):808-840.
8. Rubin DB. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Stat Med.* 2007;26(1):20-36.
9. Hill AB. The environment and disease: association or causation?. *J Royal Soc Med.* 1965;58(5):295-300.
10. Doll R, Hill AB. Mortality in relation to smoking: ten years' observations of British doctors. *Br Med J.* 1964;1(5395):1399.
11. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol.* 1974;66(5):688.
12. Rubin DB. Comment: Neyman (1923) and causal inference in experiments and observational studies. *Stat Sci.* 1990;5(4):472-480.
13. Neyman J, Iwazskiewicz K. Statistical problems in agricultural experimentation. *Suppl J Royal Stat Soc.* 1935;2(2):107-180.
14. Hahn J. On the role of the propensity score in efficient semiparametric estimation of average treatment effects. *Econometrica.* 1998;315-331.
15. Geiger D, Pearl J. On the logic of causal models. *Mach Intell Pattern Recognit.* 1990;9:3-14.
16. Zhang NL, Poole D. Exploiting causal independence in Bayesian network inference. *J Artif Intell Res.* 1996;5:301-328.
17. VanderWeele TJ, Hernán MA. From counterfactuals to sufficient component causes and vice versa. *Eur J Epidemiol.* 2006;21(12):855-858.
18. Flanders WD. On the relationship of sufficient component cause models with potential outcome (counterfactual) models. *Eur J Epidemiol.* 2006;21(12):847-853.
19. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology.* 1999;1:37-48.
20. Glymour MM. Using causal diagrams to understand common problems in social epidemiology. *Methods Soc Epidemiol.* 2006:393-428.
21. Spirtes P, Zhang K. Causal discovery and inference: concepts and recent methodological advances. *Appl Inform (Berl).* 2016;3(1):3.
22. Imbens G. Potential outcome and directed acyclic graph approaches to causality: relevance for empirical practice in economics. *Natl Bureau Econ Res.* 2019:w26104.
23. Rothman KJ, Greenland S. Causation and causal inference in epidemiology. *Am J Public Health.* 2005;95(S1):S144-S150.
24. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med.* 2010;152(11):726-732.
25. Anderson ML, Chiswell K, Peterson ED, Tasneem A, Topping J, Califf RM. Compliance with results reporting at ClinicalTrials.gov. *N Engl J Med.* 2015;372(11):1031-1039.
26. Von Elm E, Altman DG, Egger M, et al. The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. *Int J Surg.* 2014;12(12):1495-1499.
27. Holland PW. Statistics and causal inference. *J Am Stat Assoc.* 1986;81(396):945-960.
28. Westreich D, Cole SR. Invited commentary: positivity in practice. *Am J Epidemiol.* 2010;171(6):674-677.
29. Holland PW, Rubin DB. Causal inference in retrospective studies. *ETS Res Rep Ser.* 1987;1987(1):203-231.
30. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol.* 2016;183(8):758-764.
31. Hartman E, Grieve R, Ramsahai R, Sekhon JS. From sample average treatment effect to population average treatment effect on the treated: combining experimental with observational studies to estimate population treatment effects. *J Royal Stat Soc Ser A (Stat Soc).* 2015;178(3):757-778.
32. Petersen ML. Applying a causal road map in settings with time-dependent confounding. *Epidemiology.* 2014;25(6):898.
33. Petersen ML, van der Laan MJ. Causal models and learning from data: integrating causal modeling and statistical estimation. *Epidemiology.* 2014;25(3):418.
34. Pearl J. Graphs, causality, and structural equation models. *Social Methods Res.* 1998;27(2):226-284.
35. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology.* 2000;11(5):550-560.
36. Austin PC. Some methods of propensity-score matching had superior performance to others: results of an empirical investigation and Monte Carlo simulations. *Biom J.* 2009;51(1):171-184.
37. Austin PC. A comparison of 12 algorithms for matching on the propensity score. *Stat Med.* 2014;33(6):1057-1069.
38. Stuart EA. Matching methods for causal inference: a review and a look forward. *Stat Sci.* 2010;25(1):1.
39. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika.* 1983;70(1):41-55.
40. Rosenbaum PR, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. *J Am Stat Assoc.* 1984;79(387):516-524.
41. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Stat Med.* 2015;34(28):3661-3679.
42. Austin PC. The performance of different propensity-score methods for estimating relative risks. *J Clin Epidemiol.* 2008;61(6):537-545.
43. Austin PC. The performance of different propensity score methods for estimating marginal hazard ratios. *Stat Med.* 2013;32(16):2837-2849.
44. Ouyang F. Catching the falling star: points to consider when using propensity scores. *Ann Rheum Dis.* 2020;79(3):e26.
45. Stürmer T, Wang T, Golightly YM, Keil A, Lund JL, Jonsson Funk M. Methodological considerations when analysing and interpreting real-world data. *Rheumatology.* 2020;59(1):14-25.
46. Liu K, Tomlinson G, Reed AM, et al. Pilot study of the juvenile dermatomyositis consensus treatment plans: a CARRA Registry study. *J Rheumatol.* 2020. In press.
47. Yasunaga H. Introduction to Applied Statistics—Chapter 1 Propensity Score Analysis. *Ann Clin Epidemiol.* 2020;2(2):33-37.
48. Moons P. Propensity weighting: how to minimise comparative bias in non-randomised studies? *Eur J Cardiovasc Nurs.* 2020;19(1):83-88.
49. DeSantis SM, Swartz MD, Greene TJ, et al. Interim monitoring of nonrandomized prospective studies that invoke propensity scoring for decision making. *J Trauma Acute Care Surg.* 2020;88(2):e46-e52.
50. Gao L, Rosenberg MA. Assessing the causal impact of delayed oral health care on emergency department utilization. *North Am Actuarial J.* 2020. In press.
51. Bica I, Alaa AM, Lambert C, van der Schaar M. From real-world patient data to individualized treatment effects using machine learning: current and future methods to address underlying challenges. *Clin Pharmacol Ther.* 2020. In press.
52. Lu N, Xu Y, Yue LQ. Some considerations on design and analysis plan on a nonrandomized comparative study using propensity score methodology for medical device premarket evaluation. *Stat Biopharm Res.* 2020;12(2):155-163.
53. Toh S. Analytic and data sharing options in real-world multi-database studies of comparative effectiveness and safety of medical products. *Clin Pharmacol Ther.* 2020;107(4):834-842.
54. Izem R, Liao J, Hu M, et al. Comparison of propensity score methods for pre-specified subgroup analysis with survival data. *J Biopharm Stat.* 2020;3(4):734-751.
55. Gray CM, Grimson F, Layton D, Pocock S, Kim J. A framework for methodological choice and evidence assessment for studies using external comparators from real-world data. *Drug Saf.* 2020;43(7):623.
56. Li X, Lu CY. Pharmacoepidemiological approaches in health care. In: Babar ZUD, ed. *Pharmacy Practice Research Methods.* Singapore: Springer; 2020:171-202.
57. Williams G, Maroufy V, Rasmy L, et al. Vasopressor treatment and mortality following nontraumatic subarachnoid hemorrhage:

- a nationwide electronic health record analysis. *Neurosurg Focus*. 2020;48(5):E4.
58. Alhamad T, Kunjal R, Wellen J, et al. Three-month pancreas graft function significantly influences survival following simultaneous pancreas-kidney transplantation in type 2 diabetes patients. *Am J Transplant*. 2020;20(3):788-796.
  59. Elsayed ME, Morris AD, Li X, Browne LD, Stack AG. Propensity score matched mortality comparisons of peritoneal and in-centre haemodialysis: systematic review and meta-analysis. *Nephrol Dial Transplant*. 2020. In press.
  60. Steele J, Fisher J, Crawford D. Does increasing an athletes' strength improve sports performance? A critical review with suggestions to help answer this, and other, causal questions in sport science. *J Trainol*. 2020;9(1):20.
  61. Bastiaannet E. Research methods: epidemiologic research in geriatric oncology. In: Extermann M, ed. *Geriatric Oncology*. Cham, Switzerland: Springer; 2020:1031-1042.
  62. Peng L, Chen JL, Zhu GL, et al. Treatment effects of cumulative cisplatin dose during radiotherapy following induction chemotherapy in nasopharyngeal carcinoma: propensity score analyses. *Ther Adv Med Oncol*. 2020. In press.
  63. Huang WK, Chang SH, Hsu HC, et al. Postdiagnostic metformin use and survival of patients with colorectal cancer: a nationwide cohort study [published online ahead of print April 8, 2020]. *Int J Cancer*. 2020;147(7):1904-1916. doi:10.1002/ijc.32989.
  64. Gaudineau A, Lorthé E, Quere M, et al. Planned delivery route and outcomes of cephalic singletons born spontaneously at 24 to 31 weeks' gestation: the EPIPAGE-2 cohort study. *Acta Obstetr Gynecol Scand*. 2020. In press.
  65. Yu YH, Bodnar LM, Himes KP, Brooks MM, Naimi AI. Association of overweight and obesity development between pregnancies with stillbirth and infant mortality in a cohort of multiparous women. *Obstetr Gynecol*. 2020;135(3):634-643.
  66. Highland KB, Soumoff AA, Spinks EA, Kemezis PA, Buckenmaier CC III. Ketamine administration during hospitalization is not associated with posttraumatic stress disorder outcomes in military combat casualties: a matched cohort study. *Anesth Analg*. 2020;130(2):402-408.
  67. Peristera P, Platts LG, Magnusson Hanson LL, Westerlund H. A comparison of the B-spline group-based trajectory model with the polynomial group-based trajectory model for identifying trajectories of depressive symptoms around old-age retirement. *Aging Ment Health*. 2020;24(3):445-452.
  68. Azagba S, Shan L, Latham K, Qeadan F. Disparities in adult cigarette smoking and smokeless tobacco use by sexual identity. *Drug Alcohol Depend*. 2020;206:107684.
  69. Ashton RA, Prosnitz D, Andrada A, Herrera S, Yé Y. Evaluating malaria programmes in moderate-and low-transmission settings: practical ways to generate robust evidence. *Malar J*. 2020;19(1):1-4.
  70. Soejima T, Sato I, Takita J, et al. Impacts of physical late effects on presenteeism in childhood cancer survivors. *Pediatr Int*. 2020. In press.
  71. Coffman DL, Zhou J, Cai X. Comparison of methods for handling covariate missingness in propensity score estimation with a binary exposure. *BMC Med Res Methodol*. 2020;20(1):1-4.
  72. Stack CB, Meibohm AR, Liao JM, Guallar E. Studies using randomized trial data to compare nonrandomized exposures. *Ann Intern Med*. 2020;172(7):492-494.
  73. Zhu J, Gallego B. Targeted estimation of heterogeneous treatment effect in observational survival analysis. *J Biomed Inform*. 2020;107:103474.
  74. Hu L, Gu C, Lopez M, Ji J, Wisnivesky J. Estimation of causal effects of multiple treatments in observational studies with a binary outcome. *Stat Methods Med Res*. 2020. In press.
  75. Austin PC, Thomas N, Rubin DB. Covariate-adjusted survival analyses in propensity-score matched samples: imputing potential time-to-event outcomes. *Stat Methods Med Res*. 2020;29(3):728-751.
  76. Chatterjee E, Sennott C. Fertility intentions and maternal health behaviour during and after pregnancy. *Popul Stud*. 2020;74(1):55-74.
  77. Welberry HJ, Brodaty H, Hsu B, Barbieri S, Jorm LR. Impact of prior home care on length of stay in residential care for Australians with dementia. *J Am Med Dir Assoc*. 2020;21(6):843-850.e5.
  78. Guo S, Fraser MW. *Propensity Score Analysis: Statistical Methods and Applications*. Thousand Oaks, CA: SAGE Publications; 2014.
  79. Pan W, Bai H. *Propensity Score Analysis*. New York, NY: Guilford Publications; 2015.
  80. Steiner PM, Cook D. Matching and propensity scores. The Oxford handbook of quantitative methods. 2013;1:237-259.
  81. Holmes WM. *Using Propensity Scores in Quasi-Experimental Designs*. Thousand Oaks, CA: Sage Publications; 2013.
  82. Leite W. *Practical Propensity Score Methods Using R*. Thousand Oaks, CA: Sage Publications; 2016.
  83. Forbes SP, Dahabreh IJ. Benchmarking observational analyses against randomized trials: a review of studies assessing propensity score methods. *J Gen Intern Med*. 2020;35(5):1396-1404.
  84. Shah BR, Laupacis A, Hux JE, Austin PC. Propensity score methods gave similar results to traditional regression modeling in observational studies: a systematic review. *J Clin Epidemiol*. 2005;58(6):550-559.
  85. Gayat E, Pirracchio R, Resche-Rigon M, Mebazaa A, Mary JY, Porcher R. Propensity scores in intensive care and anaesthesia literature: a systematic review. *Intensive Care Med*. 2010;36(12):1993-2003.
  86. Thoemmes FJ, Kim ES. A systematic review of propensity score methods in the social sciences. *Multivariate Behav Res*. 2011;46(1):90-118.
  87. Weitzen S, Lapane KL, Toledano AY, Hume AL, Mor V. Principles for modeling propensity scores in medical research: a systematic literature review. *Pharmacoepidemiol Drug Saf*. 2004;13(12):841-853.
  88. Austin PC. A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003. *Stat Med*. 2008;27(12):2037-2049.
  89. Zakrisson TL, Austin PC, McCredie VA. A systematic review of propensity score methods in the acute care surgery literature: avoiding the pitfalls and proposing a set of reporting guidelines. *Eur J Trauma Emerg Surg*. 2018;44(3):385-395.
  90. Funk MJ, Westreich D, Wiessen C, Stürmer T, Brookhart MA, Davidian M. Doubly robust estimation of causal effects. *Am J Epidemiol*. 2011;173(7):761-767.
  91. Austin PC. The performance of different propensity-score methods for estimating differences in proportions (risk differences or absolute risk reductions) in observational studies. *Stat Med*. 2010;29(20):2137-2148.
  92. Imai K, Ratkovic M. Covariate balancing propensity score. *J Royal Soc Ser B Stat Methodol*. 2014;76(1):243-263.
  93. Pearl J. Causal diagrams for empirical research. *Biometrika*. 1995;82(4):669-688.
  94. Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. *Stat Med*. 2014;33(13):2297-2340.
  95. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol*. 2000;29(4):722-729.
  96. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *J Am Stat Assoc*. 1996;91(434):444-455.
  97. Angrist JD, Imbens GW. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *J Am Stat Assoc*. 1995;90(430):431-442.
  98. Hearst N, Newman TB, Hulley SB. Delayed effects of the military draft on mortality. *N Engl J Med*. 1986;314(10):620-624.
  99. Hearst N, Newman TB. Proving cause and effect in traumatic stress: the draft lottery as a natural experiment. *J Trauma Stress*. 1988;1(2):173-180.
  100. Mansournia MA, Etminan M, Danaei G, Kaufman JS, Collins G. Handling time varying confounding in observational research. *BMJ*. 2017;359:j4587.
  101. Naimi AI, Cole SR, Kennedy EH. An introduction to g methods. *Int J Epidemiol*. 2017;46(2):756-762.
  102. Daniel RM, Cousens SN, De Stavola BL, Kenward MG, Sterne JA. Methods for dealing with time-dependent confounding. *Stat Med*. 2013;32(9):1584-1618.
  103. Robins J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the

- healthy worker survivor effect. *Math Model.* 1986;7(9-12):1393-1512.
104. Taubman SL, Robins JM, Mittleman MA, Hernán MA. Intervening on risk factors for coronary heart disease: an application of the parametric g-formula. *Int J Epidemiol.* 2009;38(6):1599-1611.
  105. Hernán MA, Robins JM. Estimating causal effects from epidemiological data. *J Epidemiol Community Health.* 2006;60(7):578-586.
  106. Robins JM, Hernán MA. Estimation of the causal effects of time-varying exposures. *Longitudinal Data Anal.* 2009;553:599.
  107. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol.* 2008;168(6):656-664.
  108. Hogan JW, Lancaster T. Instrumental variables and inverse probability weighting for causal inference from longitudinal observational studies. *Stat Methods Med Res.* 2004;13(1):17-48.
  109. van der Laan MJ, Gruber S. Collaborative double robust targeted maximum likelihood estimation. *Int J Biostat.* 2010;6(1):17.
  110. Schuler MS, Rose S. Targeted maximum likelihood estimation for causal inference in observational studies. *Am J Epidemiol.* 2017;185(1):65-73.
  111. Xu Y, Xu Y, Saria S. A Bayesian nonparametric approach for estimating individualized treatment-response curves. *Mach Learn Healthc Conf.* 2016;56:282-300.
  112. Rubin DB. *Matched Sampling for Causal Effects.* Cambridge, England: Cambridge University Press; 2006.
  113. VanderWeele T. *Explanation in Causal Inference: Methods for Mediation and Interaction.* Oxford, England: Oxford University Press; 2015.
  114. Van der Laan MJ, Rose S. *Targeted Learning: Causal Inference for Observational and Experimental Data.* Berlin, Germany: Springer Science & Business Media; 2011.
  115. Guyon I, Statnikov A, Batu BB, eds. *Cause Effect Pairs in Machine Learning.* Cham, Switzerland: Springer; 2019.
  116. Pearl J. *Causality.* Cambridge, England: Cambridge University Press; 2009.
  117. Pearl J, Mackenzie D. *The Book of Why: The New Science of Cause and Effect.* New York, NY: Basic Books; 2018.
  118. Pearl J, Glymour M, Jewell NP. *Causal Inference in Statistics: A primer.* Hoboken, NJ: John Wiley & Sons; 2016.
  119. Shipley B. *Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference With R.* Cambridge, England: Cambridge University Press; 2016.
  120. Mulaik SA. *Linear Causal Modeling With Structural Equations.* Boca Raton, FL: CRC Press; 2009.
  121. Halpern JY. *Actual Causality.* Cambridge, MA: The MIT Press; 2016.
  122. Peters J, Janzing D, Schölkopf B. Cambridge, MA: *Elements of Causal Inference.* The MIT Press; 2017.
  123. Hernán MA, Robins JM. *Causal Inference: What if.* Boca Raton, FL: Chapman & Hill/CRC. 2020.
  124. Rosenbaum PR. *Observation and Experiment.* Cambridge, MA: Harvard University Press; 2017.
  125. Berzuini C, Dawid P, Bernardinell L, eds. *Causality: Statistical Perspectives and Applications.* Hoboken, NJ: John Wiley & Sons; 2012.
  126. Morgan SL, Winship C. *Counterfactuals and Causal Inference.* Cambridge, England: Cambridge University Press; 2015.
  127. Aickin M. *Causal Analysis in Biomedicine and Epidemiology: Based on Minimal Sufficient Causation.* Boca Raton, FL: CRC Press; 2001.
  128. Imbens GW, Rubin DB. *Causal Inference in Statistics, Social, and Biomedical Sciences.* Cambridge, England: Cambridge University Press; 2015.
  129. Salmon WC. *Causality and Explanation.* Oxford, England: Oxford University Press; 1998.
  130. Gelman A, Meng XL, eds. *Applied Bayesian Modeling and Causal Inference From Incomplete-Data Perspectives.* Hoboken, NJ: John Wiley & Sons; 2004.
  131. Rohlfing I. *Case Studies and Causal Inference: An Integrative Framework.* London, England: Palgrave Macmillan; 2012.
  132. Morgan SL, ed. *Handbook of Causal Analysis for Social Research.* New York, NY: Springer; 2013.
  133. He H, Wu P, Chen DG, eds. *Statistical Causal Inferences and Their Applications in Public Health Research.* Cham, Switzerland: Springer International Publishing; 2016.
  134. Murnane RJ, Willett JB. *Methods Matter: Improving Causal Inference in Educational and Social Science Research.* Oxford, England: Oxford University Press; 2010.
  135. Freedman DA. *Statistical Models and Causal Inference: A Dialogue With the Social Sciences.* Cambridge, England: Cambridge University Press; 2010.
  136. Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology.* Philadelphia, PA: Lippincott Williams & Wilkins; 2008.
  137. Westreich D. *Epidemiology by Design: A Causal Approach to the Health Sciences.* Oxford, England: Oxford University Press; 2019.
  138. Oakes JM, Kaufman JS, eds. *Methods in Social Epidemiology.* Hoboken, NJ: John Wiley & Sons; 2017.
  139. Gatsonis C, Morton SC, eds. *Methods in Comparative Effectiveness Research.* Boca Raton, FL: CRC Press; 2017.
  140. Shadish WR, Cook TD, Campbell DT. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Boston, MA: Houghton Mifflin; 2002.
  141. Xiong M. *Big Data in Omics and Imaging: Integrated Analysis and Causal Inference.* Boca Raton, FL: CRC Press; 2018.
  142. Goertzel B, Geisweiller N, Coelho L, Janičić P, Pennachin C. *Real-World Reasoning: Toward Scalable, Uncertain Spatiotemporal, Contextual and Causal Inference.* Berlin, Germany: Springer Science & Business Media; 2011.
  143. Best H, Wolf C, eds. *The SAGE Handbook of Regression Analysis and Causal Inference.* Thousand Oaks, CA: Sage; 2014.
  144. <https://sites.google.com/view/ocis/>. Accessed July 4, 2020.
  145. <https://www.coursera.org/search?query=causal%20inference&>. Accessed July 4, 2020.
  146. [projects.illc.uva.nl/cil/](https://projects.illc.uva.nl/cil/). Accessed July 4, 2020.
  147. [bayes.cs.ucla.edu/jp\\_home.html](https://bayes.cs.ucla.edu/jp_home.html). Accessed July 4, 2020.
  148. [www.cceb.med.upenn.edu/cci](https://www.cceb.med.upenn.edu/cci). Accessed July 4, 2020.
  149. [www.hsph.harvard.edu/causal/software/](https://www.hsph.harvard.edu/causal/software/). Accessed July 4, 2020.
  150. <http://www.biostat.jhsph.edu/~estuart/propensityscoresoftware.html>. Accessed July 4, 2020.
  151. [www.healthpolicyinstitute.pitt.edu/cerc](https://www.healthpolicyinstitute.pitt.edu/cerc). Accessed July 4, 2020.
  152. [www.landsittellab.pitt.edu/educational-resources](https://www.landsittellab.pitt.edu/educational-resources). Accessed July 4, 2020.
  153. [cerbot.org/](https://cerbot.org/). Accessed July 4, 2020.
  154. [www.pcori.org/research-results?f%5B0%5D=field\\_project\\_type%3A298&f%5B1%5D=field\\_award\\_research\\_priority%3A161#search-results](https://www.pcori.org/research-results?f%5B0%5D=field_project_type%3A298&f%5B1%5D=field_award_research_priority%3A161#search-results). Accessed July 4, 2020.
  155. [www.pcori.org/research-results/about-our-research/research-methodology](https://www.pcori.org/research-results/about-our-research/research-methodology). Accessed July 4, 2020.
  156. [www.landsittellab.pitt.edu/decode-cer-tool](https://www.landsittellab.pitt.edu/decode-cer-tool). Accessed July 4, 2020.
  157. [www.landsittellab.pitt.edu/educational-resources/online-self-guided-course-propensity-score-based-methods](https://www.landsittellab.pitt.edu/educational-resources/online-self-guided-course-propensity-score-based-methods). Accessed July 4, 2020.
  158. [www.ccd.pitt.edu](https://www.ccd.pitt.edu). Accessed July 5, 2020.