SHEA WHITE PAPER

Research Methods in Healthcare Epidemiology and Antimicrobial Stewardship—Observational Studies

Graham M. Snyder, MD, SM;[1] Heather Young, MD;[2] Meera Varman, MD;[3] Aaron M. Milstone, MD, MHS;[4] Anthony D. Harris, MD, MPH, FIDSA, FSHEA;[5] Silvia Munoz-Price, MD, PhD[6]

Observational studies compare outcomes among subjects with and without an exposure of interest, without intervention from study investigators. Observational studies can be designed as a prospective or retrospective cohort study or as a case-control study. In healthcare epidemiology, these observational studies often take advantage of existing healthcare databases, making them more cost-effective than clinical trials and allowing analyses of rare outcomes. This paper addresses the importance of selecting a well-defined study population, highlights key considerations for study design, and offers potential solutions including biostatistical tools that are applicable to observational study designs.

Infect Control Hosp Epidemiol 2016;37:1141–1146

BACKGROUND

Observational studies are a common study design in healthcare epidemiology, in large part due to the increasing accessibility of electronic data to clinicians, infection preventionists, and administrators. In contrast to randomized controlled trials, the investigator does not intervene with the exposure, and generally, the resources required for the study are far less intensive. Observational studies provide an opportunity to define and elucidate potential cause-and-effect relationships when it is not feasible to perform a randomized controlled trial. In this review, we discuss strategies for designing and conducting observational studies to maximize reliability and validity. We encourage readers to review the *Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement*, or any of several other helpful tools, when planning an observational study.[1,2] Best practices in reporting your study begin with careful attention to the design of the study.

DESIGNING AN OBSERVATIONAL STUDY

Like any investigation attempting to explore the relationship between an exposure and potential outcome, the objective of most observational studies is to test the counterfactual, ie, what the outcome would have been if the subject had not received the exposure of interest.[3] To do this, all observational studies start with a cohort (ie, the population, or a specified subset of the population) at risk for the outcome of interest and then compare the likelihood of the outcome among subjects with and without the exposure of interest.

To this end, observational studies may have a wide range of designs, including retrospective cohort, prospective cohort, and case-control studies. In cohort studies, all individuals in the defined population at risk for the outcome are followed for a specified period of time to ascertain these events. Subjects in the cohort may be defined by a shared characteristic, location, and/or time period (eg, all residents of a skilled nursing facility during the 2010–2011 influenza season). All individuals at risk for the outcome are included in the analysis. Retrospective cohorts include a study population for whom the outcome of interest has occurred at the time of study design and enrollment; in contrast, subjects in a prospective cohort are enrolled before the occurrence of outcomes. The choice of retrospective versus prospective design is often dictated by the circumstances under which the study is performed. Prospective cohort studies may afford the investigator an opportunity to improve the completeness and reliability of the data collected, though the time to complete the study may be longer. Although some studies may involve both retrospective and prospective data collection,[4] within the realm of healthcare epidemiology, a completely retrospective design for cohort studies is often most convenient to conduct. Debate sometimes ensues regarding the definitions of "retrospective" and "prospective" observational studies. A distinction between timing and directionality may be made: the timing of a study is the relationship between the occurrence of outcomes and data

Affiliations: 1. Beth Israel Deaconess Medical Center, Harvard University, Boston, Massachusetts; 2. Denver Health Medical Center, University of Colorado Hospital, Denver, Colorado; 3. Creighton University School of Medicine, Omaha, Nebraska; 4. Johns Hopkins Medical Institutions, Baltimore, Maryland; 5. University of Maryland School of Medicine, Veterans Affairs Maryland Health Care System, Baltimore, Maryland; 6. Institute for Health and Society and Department of Medicine, Medical College of Wisconsin, Milwaukee, Wisconsin.

Received April 25, 2016; accepted April 26, 2016; electronically published June 20, 2016

 0899-823X/2016/3710-0002. DOI: 10.1017/ice.2016.118

collection (retrospective vs prospective), whereas the directionality of the study refers to the order in which the exposure and outcome are identified in the study cohort (forward vs backward). In a case-control study, individuals who developed the outcome (cases) and did not develop the outcome (controls) are first identified from the same base population as an investigator would define the source population for a cohort study. There are various ways to select controls, but regardless of the selection method, it is critical to ensure that controls are drawn from the same source population as the cases to help mitigate biases.[5–7] After appropriately defining the cohort and selecting cases and controls, the association between the exposure of interest and the outcome may be analyzed.

ADVANTAGES AND DISADVANTAGES

Observational studies in healthcare epidemiology benefit from the widespread use of electronic data collection (Table 1). Data collected for clinical or administrative purposes may come with potential limitations.[8,12] However, the availability of these data, the relatively low cost of conducting an observational study, and the ability to evaluate infrequent outcomes make cohort and case-control studies the most commonly implemented study designs in the field.

Observational studies are highly susceptible to commonly encountered biases such as selection, assessment, time-dependent, loss to follow-up, and recall biases. These biases may be introduced by a mis-specified or poorly defined source population, use of a non-standard or subjective definition of the outcome, difficulty in ascertaining covariates that impact the outcome of interest, incomplete accounting of time at risk, and confounding from both measured and unmeasured confounders.[9,12]

PITFALLS AND TIPS

A "perfect" observational study is difficult, if not impossible, to achieve due to the intrinsic nature of observational study designs. For this reason, investigators should aim to design and conduct a study with the fewest limitations feasible, and readers should carefully consider the impact of study limitations while evaluating the conclusions of the study. Table 1 enumerates potential pitfalls to avoid and tips to resolve these limitations.

It is important to start with a clearly defined study premise: does the study define a discrete hypothesis that is testable with the data available (or to be collected), and does the study address a meaningful question not sufficiently answered in the existing peer-reviewed published literature? While a cohort or case-control study may be chosen for cost and convenience and is often the appropriate design for healthcare epidemiology research, alternative designs such as time-series analysis or controlled trial may be more appropriate for some hypotheses and conditions.

As discussed in the introduction, selection of cases and controls is an important and underappreciated consideration. Defining the outcome and consideration of covariates (measurable and unmeasurable) are important to ensure validity and avoid potential biases, including confounding. Outcomes in healthcare epidemiology may be a measure of incidence or prevalence, with or without consideration of recurrence. As prevalence is influenced not only by incidence but also by duration of disease, it is preferable to use incidence as an outcome in most circumstances, though the data collection and analysis may be more complex.[13]

The precise and accurate collection of data is an essential aspect of internal validity (how well the study was designed and performed). If multiple reviewers collect data, a calculation of the inter-rater reliability may be used to assess the precision and internal validity of the study.[14,15] External validity (the generalizability of the study findings to other populations) can be limited if variables or the cohort are poorly defined. Standard objective definitions of exposures and other variables, but particularly the outcome, allow for optimal interpretation of the study findings. In hospital epidemiology, a number of resources for standard surveillance definitions are available, such as National Healthcare Safety Network guidance and criteria for identifying and defining healthcare-associated infections.[16] Table 2 provides a condensed checklist of key considerations that may help identify and address potential pitfalls in the design phase of a study.

STATISTICAL CONSIDERATIONS

Because the "assignment" of exposures and potential confounders to patients in the study population is beyond the control of the investigator conducting an observational study, statistical tools are often employed to account for associations that may affect the measured relationship between the exposure and the outcome. These analyses may be complex and require the assistance of a biostatistician, including both before initiating a study and during the analysis phase. A biostatistician and/or an epidemiologist trained in study design may advise on complex situations: recurrent outcomes per unique subject (eg, recurrent *Clostridium difficile* colitis); nonbinary risk categories (eg, the risk of healthcare associated infection due to *Staphylococcus aureus* among patients with a baseline history of infection, a positive colonization screen, a negative screen, and/or unknown colonization status); or exposure variables that are time dependent (eg, determining association between the duration of antimicrobial exposure and risk of infection due to carbapenem-resistant Enterobacteriaceae). In Table 3, we have outlined examples some of the fundamental tools used in the analysis of observational studies.

OBSERVATIONAL STUDIES IN HEALTHCARE EPIDEMIOLOGY AND ANTIMICROBIAL STEWARDSHIP

While there are no limitation-free observational studies, several examples will help illustrate key principles of this review.

https://doi.org/10.1017/ice.2016.118 Published online by Cambridge University Press

TABLE 1. Advantages, Disadvantages, and Potential Pitfalls of Using Observational Studies in Healthcare Epidemiology and Antimicrobial Stewardship Research

Advantages	Notes
"Natural" experiments may allow studies not feasible through clinical trials.	"Natural" experiments ideally include assignment to an exposure group for reasons unrelated to potential confounders.[21]
Data are often readily available in existing sources.	The expansion of electronic medical records has improved data accessibility and ease of use.
Lower costs compared with a randomized controlled trial with similar exposure and outcome of interest	Costs of data acquisition and review may still be substantial.
Allows for evaluation of low-frequency outcomes	Case-control study may improve efficiency without significant loss in power.
Allows for flexibility based upon practical variables	Retrospective or prospective designs may be chosen based upon available resources.
Cohort design allows for calculation of risk ratio	Risk ratio is more intuitive than odds ratio or hazard ratio (generally required for case-control studies).[22]
Disadvantages	**Notes**
Data collection may be time-consuming if manual review is required.	Measurement of variables including exposure and outcome may be beyond the control of investigators.
Existing datasets may include substantial missing or erroneous data.	Various methods are available to partially accommodate missing data.[23]
Unmeasured variables may confound the relationship between exposure and outcome.	Not all potential confounders may be measurable; some advanced statistical methods may attempt to quantify unmeasured confounding, or the study may be conducted through a randomized design.
A significant relationship may only support a statistical association rather than demonstrating causality.	Support for causality requires biological plausibility, consistency, previous evaluations of alternative explanations and potentially other supporting evidence.
Prone to type I errors (detecting an association between exposure and outcome that does not exist)	These errors may include chance associations (particularly when there are a large number of associations tested) or reflect 1 or more biases.
Potential pitfalls	**Tips and solutions**
Lack of a discrete aim and hypothesis, including defined exposure and outcomes	"Exploratory" studies may be appropriate if judiciously planned prior to conducting the study.
Aim does not address a problem of interest to the community.	Ensures that the research question has not already been sufficiently reported in the literature, or addressed by a substantially more rigorous study (eg, a randomized controlled trial)
Multiple concomitant interventions or a composite outcome may limit interpretability.	Multiple interventions may be appropriate to undertake (eg, response to an outbreak); if sequential, analytic methods such as time-series analysis may be useful.
Cases and controls are not drawn from the same source population.	Clearly define a population for which cases and controls may be present and at risk of the outcome
Cases and controls are not at risk for the outcome for the follow-up duration, or for the same amount of time.	Ensure that for the period of time that both cases and controls are considered for the outcome, they are truly at risk for the outcome. Choose a time period from exposure to outcome that is biologically plausible. Consider carefully defining the outcome as a prevalent, incident, or recurrent event.
Use of nonstandard outcome definitions, or inadequate outcome ascertainment	Must be reliable (eg, assessed through inter-rater reliability) and valid (outcome measures what is intended) Consider pros/cons of clinically defined versus surveillance definitions
Use of surrogate (or proxy) measures rather than the event of interest	The defined outcome should be a meaningful end point, or use a surrogate outcome that is proven to predict the end point of interest (e.g., defining nosocomial acquisition as colonization versus infection).
Bias due to unmeasured confounding variables	Consider and quantify to the extent feasible all variables that may be confounders, based on proven or theoretical relationship. Consider possible unmeasured confounders and their potential impact.
Insufficient power, or type 2 error (failing to detect a difference when one exists)	Perform a power or sample size calculation based on published data to plan a study size adequate for the frequency of the outcome
Potential statistical complexity	Consider employing the assistance of a biostatistician, during the planning stages of the study

TABLE 2. Checklist of Key Considerations When Developing an Observational Study in Healthcare Epidemiology

1. What is the research question you hope to answer? (Establish the study hypothesis prior to initiation of the study.)
2. Which type of observational study best addresses the research question and accommodates the data available or to be collected?
3. What type of participants or groups will be studied? (Clearly define and identify the type of individual subjects or groups that will be included. Specify inclusion and exclusion criteria.)
4. How many subjects will be required? (Power or sample-size calculations based on existing data will help avoid a type 2 error.)
5. What outcomes will be evaluated? Is the outcome meaningful, generalizable, and does it reflect a measure that will improve patient care?
6. How will an association between the exposure and outcome be analyzed?
7. What potential confounding variables should be addressed? (Develop a statistical analysis plan prior to collection of outcome data. The study plan should include consideration of the time at risk for the outcome and potential confounders, including those that are measurable and unmeasurable.)
8. What important biases is the study design susceptible to? (Consider the important biases and ensure that the biases are not large enough to negate the value of performing the study.)
9. For case-control studies, have the controls been selected from the base population of interest? (Controls should be subjects who could be considered among the same group of individuals as the cases, and at similar risk of the outcome.)
10. For cohort studies, is the study design able to identify subjects lost to follow-up? (Ensure that loss to follow up does not negate the value of performing the study and try to minimize loss to follow-up.)
11. Will you be able to report your study design and findings completely? (Consider using a systematic checklist.[1])

For case-control studies, appropriately identifying cases and controls is essential to testing the counterfactual: did every subject have the same opportunity to be exposed to risk factors of interest and to develop the outcome? In a study that aimed to investigate risk factors for the acquisition of extended-spectrum β-lactamase–producing *Klebsiella pneumoniae* in an intensive care unit (ICU), cases were defined as patients for whom the organism was isolated during the ICU admission, and controls were selected from the same ICU population with a similar length of exposure in the ICU (and who were not defined as a case), and were matched on age, gender, severity of illness score, and underlying disease.[17] Having defined cases and controls in this fashion—including consideration of control group selection and adjustment for time at risk— allowed the investigators to be more confident that subjects who did not develop the outcome had the potential to become a case had their exposures been different.[18]

In prospective cohort studies, by controlling which data elements or specimens to collect, investigators may have the opportunity to improve the internal validity and overall quality of the study. In a 2011 study, Dutch investigators included in their prospective cohort consecutive patients admitted to 1 of 4 ICUs for at least 48 hours during the study period. Investigators collected surveillance cultures from the respiratory tract (on admission, twice weekly, and on discharge) to identify patients (cases) who acquired a multidrug-resistant *Pseudomonas aeruginosa* or *Enterobacter* sp. With this design, the investigators were able to confidently identify subjects in the cohort at risk for colonization with a resistant strain, and they had full case ascertainment in investigating variables associated with the development of resistance.[19]

Using a large, validated, multicenter, quality improvement database, investigators in California studied a cohort of 20,934 very-low-birth-weight infants born between 2002 and 2006 in a study hospital.[4] The exposure of interest was birth in a hospital participating in a quality improvement project (vs not participating in the project), and the outcome was nosocomial infection events. The outcome was carefully defined using a standardized definition and a defined at-risk period (because nosocomial infection could not occur within 4 days of birth by definition, those with length of stay <4 days were excluded). Authors accounted for the potential impact of mis-specified outcomes by excluding infants with abdominal surgery or necrotizing enterocolitis (because nosocomial bloodstream infection may be confounded by these events). The investigators acknowledged the challenges of handling missing data and determining exposure when patients were transferred between hospitals.[20]

While there exists no "perfect" study, in every study it is important to design the study to minimize factors that impact the quality of the study and to acknowledge not only the limitations present but also the significance of the impact these limitations may have on the validity and the interpretation of the results.

MAJOR TAKE-HOME POINTS

Observational studies are frequently employed in healthcare epidemiology research due to the increasing availability of electronic databases, which may be more accessible to the field due to feasibility limitations of randomized controlled trials. An observational study may be designed as a prospective or retrospective cohort study, or as a case-control study, but a unifying characteristic of these designs is the comparison of an outcome among subjects exposed and unexposed to a variable of interest without intervention from study investigators.

https://doi.org/10.1017/ice.2016.118 Published online by Cambridge University Press

TABLE 3. Statistical Tools Useful for the Analysis of Observational Studies

Statistical Tool	Key Points
χ^2 or Fisher's exact test	Allows comparison of proportions. Useful in case-control studies and to determine baseline differences in cohorts. Example: Establishing the difference in proportions of *Clostridium difficile* colitis among surgical and medical patients.
T-test Mann-Whitney U test (Wilcoxon rank-sum test)	Allows comparison of the distribution of a continuous variable among 2 different groups (referenced to the mean for T-test, median for Wilcoxon rank-sum test). Useful in case-control studies and to determine baseline differences in cohorts. Example: Determining the difference in ages (means) between intensive care unit (ICU) A patients and ICU B patients.
Linear regression	Estimates the association between one or more exposure and a continuous outcome. May also be used to estimate and account for the effects of multiple variables concomitantly. Infrequently used in hospital epidemiology. Example: Evaluating the association between the number of patients in unit A and the bacterial growth (in colony-forming units) on the unit's nursing station.
Logistic regression	Estimates the association between 1 or more exposure and binary (yes or no) outcome. Allows for considering the concomitant effects of multiple variables at the same time. Widely used in hospital epidemiology given that the outcomes observed tend to be binary. Example: Estimate the association between age, gender, and/or smoking status with the development of surgical-site infections.
Survival analysis/Cox proportional hazards	Estimates the association between 1 or more variables and the occurrence of an outcome while accounting for the amount of time to the outcome or end of follow-up.[24] Allows the incorporation of the amount of time a subject is at risk, and accounts for censoring (patients lost to follow-up). Allows for the consideration of concomitant effects of multiple variables at the same time. Example: Determine the association between time to development of ventilator-associated pneumonia and proton pump inhibitors during ICU stay.
Time dependent variables (Time-varying covariate)	Accounts for timing of the exposure and change in a subject's assignment to an exposure during follow-up, within a survival analysis framework. Example: Including renal replacement therapy (which may start or stop during follow-up) among factors that may be associated with nosocomial infections during ICU stay.
Multistate models	A model that moves subjects from one state to another, including in a bidirectional fashion.[25] Contrasts the hazards of competing risks. Example: Outcome states for a multistate model may include discharged without an infection, discharged with an unresolved infection, and death.
Propensity score	A specific logistic regression model that may be used in several ways to take into account multiple variables.[26] Two common methods of using propensity scores: • Develop a model that predicts the exposure, matches exposed and unexposed subjects on similar scores, and determines the relationship between exposure and outcome after matching. • Develop a model that predicts the exposure and stratifies and analyzes exposed and unexposed subjects in groups (eg, deciles) of propensity scores.

Researchers contemplating an observational trial in hospital epidemiology should carefully consider the design of the study before collecting and analyzing data, and researchers should be particularly vigilant when defining the study cohort at risk of the outcome. Sources of bias should be considered first in the design stage, and various statistical tools are available to aid in the analysis to accommodate bias resulting from a lack of randomized comparison groups.

CONCLUSIONS

Cohort and case-control studies represent the observational designs most frequently employed in healthcare epidemiological research. Although data might be readily available and the analysis of exposures and outcomes might seem straightforward, these studies are subject to many sources of bias. Limitations of these observational studies can be addressed by careful planning, definition of subjects at risk for the outcome, and inclusion of patients with and without the exposure of interest.

ACKNOWLEDGEMENTS

Financial support: No financial support was provided relevant to this article.

Potential conflicts of interest: L.S.M.-P. reports that she has served as a speaker for Ecolab and Xenex and as a consultant for Xenex and Clorox. All other authors report no conflicts of interest relevant to this article.

Address correspondence to Graham Snyder, Mailstop SL-435, 330 Brookline Ave, Boston, MA 02215 (gsnyder@bidmc.harvard.edu).

https://doi.org/10.1017/ice.2016.118 Published online by Cambridge University Press

REFERENCES

1. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007;147:573–577.
2. Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol* 2007;36:666–676.
3. Hofler M. Causal inference based on counterfactuals. *BMC Med Res Methodol* 2005;5:28.
4. Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods*. Belmont, CA: Lifetime Learning Publications; 1982.
5. Schulz KF, Grimes DA. Case-control studies: research in reverse. *Lancet* 2002;359:431–434.
6. Wacholder S, McLaughlin JK, Silverman DT, Mandel JS. Selection of controls in case-control studies. I. Principles. *Am J Epidemiol* 1992;135:1019–1028.
7. Wacholder S, Silverman DT, McLaughlin JK, Mandel JS. Selection of controls in case-control studies. III. Design options. *Am J Epidemiol* 1992;135:1042–1050.
8. Wyllie D, Davies J. Role of data warehousing in healthcare epidemiology. *J Hosp Infect* 2015;89:267–270.
9. Harris AD, Carmeli Y, Samore MH, Kaye KS, Perencevich E. Impact of severity of illness bias and control group misclassification bias in case-control studies of antimicrobial-resistant organisms. *Infect Control Hosp Epidemiol* 2005;26:342–345.
10. Dhar S, Tansek R, Toftey EA, et al. Observer bias in hand hygiene compliance reporting. *Infect Control Hosp Epidemiol* 2010;31:869–870.
11. Nelson RE, Nelson SD, Khader K, et al. The magnitude of time-dependent bias in the estimation of excess length of stay attributable to healthcare-associated infections. *Infect Control Hosp Epidemiol* 2015;36:1089–1094.
12. Schweizer ML, Braun BI, Milstone AM. Research methods in healthcare epidemiology and antimicrobial stewardship—quasi-experimental designs. *Infect Control Hosp Epidemiol* DOI: http://dx.doi.org/10.1017/ice.2016.91. Published online April 14, 2016.
13. Rhame FS, Sudderth WD. Incidence and prevalence as used in the analysis of the occurrence of nosocomial infections. *Am J Epidemiol* 1981;113:1–11.
14. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. *Fam Med* 2005;37:360–363.
15. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol* 2012;8:23–34.
16. 2014 NHSN Data Quality Guidance and Toolkit for Reporting Facilities: Internal Validation of NHSN Patient Safety Component Data. Centers for Disease Control and Prevention website. http://www.cdc.gov/nhsn/validation/index.html. Published 2014. Accessed January 29, 2016.
17. Piroth L, Aube H, Doise JM, Vincent-Martin M. Spread of extended-spectrum beta-lactamase-producing *Klebsiella pneumoniae*: are beta-lactamase inhibitors of therapeutic value? *Clin Infect Dis* 1998;27:76–80.
18. Harris AD, Karchmer TB, Carmeli Y, Samore MH. Methodological principles of case-control studies that analyzed risk factors for antibiotic resistance: a systematic review. *Clin Infect Dis* 2001;32:1055–1061.
19. Ong DS, Jongerden IP, Buiting AG, et al. Antibiotic exposure and resistance development in *Pseudomonas aeruginosa* and *Enterobacter* species in intensive care units. *Crit Care Med* 2011;39:2458–2463.
20. Wirtschafter DD, Powers RJ, Pettit JS, et al. Nosocomial infection reduction in VLBW infants with a statewide quality-improvement model. *Pediatrics* 2011;127:419–426.
21. Petticrew M, Cummins S, Ferrell C, et al. Natural experiments: an underused tool for public health? *Public Health* 2005;119:751–757.
22. Davies HT, Crombie IK, Tavakoli M. When can odds ratios mislead? *BMJ* 1998;316:989–991.
23. Newgard CD, Lewis RJ. Missing data: how to best account for what is not known. *JAMA* 2015;314:940–941.
24. Muscedere JG, Day A, Heyland DK. Mortality, attributable mortality, and clinical events as end points for clinical trials of ventilator-associated pneumonia and hospital-acquired pneumonia. *Clin Infect Dis* 2010;51(Suppl 1):S120–S125.
25. Schumacher M, Allignol A, Beyersmann J, Binder N, Wolkewitz M. Hospital-acquired infections—appropriate statistical treatment is urgently needed! *Int J Epidemiol* 2013;42:1502–1508.
26. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 2011;46:399–424.