A Framework for Descriptive Epidemiology

Catherine R. Lesko*, Matthew P. Fox, and Jessie K. Edwards

* Correspondence to Dr. Catherine R. Lesko, Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe Street, Baltimore, MD 21205 (e-mail: clesko2@jhu.edu).

Initially submitted February 20, 2022; accepted for publication June 16, 2022.

In this paper, we propose a framework for thinking through the design and conduct of descriptive epidemiologic studies. A well-defined descriptive question aims to quantify and characterize some feature of the health of a population and must clearly state: 1) the target population, characterized by person and place, and anchored in time; 2) the outcome, event, or health state or characteristic; and 3) the measure of occurrence that will be used to summarize the outcome (e.g., incidence, prevalence, average time to event, etc.). Additionally, 4) any auxiliary variables will be prespecified and their roles as stratification factors (to characterize the outcome distribution) or nuisance variables (to be standardized over) will be stated. We illustrate application of this framework to describe the prevalence of viral suppression on December 31, 2019, among people living with human immunodeficiency virus (HIV) who had been linked to HIV care in the United States. Application of this framework highlights biases that may arise from missing data, especially 1) differences between the target population and the analytical sample; 2) measurement error; 3) competing events, late entries, loss to follow-up, and inappropriate interpretation of the chosen measure of outcome occurrence; and 4) inappropriate adjustment.

bias; checklist; data analysis; description; framework

Abbreviations: HIV, human immunodeficiency virus; NA-ACCORD, North American AIDS Cohort Collaboration on Research and Design.

Editor's note: *An invited commentary on this article appears on page 2071, and the authors' response appears on page 2073.*

Epidemiologic questions arguably exist on a continuum from purely descriptive to purely causal. To be concise, we ignore prediction questions here. There are several frameworks intended to help guide causal analyses (1, 2), but the literature on theoretical and practical guidance for conducting descriptive analyses is limited. Here we present a framework for conducting descriptive epidemiologic studies. Many, if not all, of the considerations discussed in this framework apply to estimation of valid causal effects in a population, although they may be frequently overlooked. Where there may be differences in analytical decisions depending on the type of study question, we highlight them. We summarize guidance provided herein in Table 1 in the form of a checklist modeled after the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines (3).

We define a descriptive epidemiologic question as one that aims to quantify some feature of the health of a population and, often, to characterize the distribution of that feature across the population. The estimand for causal analyses is a contrast of potential outcomes in a single population, where the potential outcomes are those we would expect to observe under some hypothetical intervention (1, 4–7). The fundamental problem of causal inference is that we cannot observe all of these potential outcomes (8). The estimand for descriptive analyses is a function of the outcomes that occurred for everyone in the target population. The estimation challenge for descriptive analyses is that we may not completely observe all of the actual outcomes. A descriptive analysis might be cross-sectional or longitudinal; it might concern a dichotomous, categorical, or continuous outcome; and it might attempt to summarize the outcome in any number of ways (e.g., median time to some event,

Table 1. Items That Should Be Included in Reports of Descriptive Studies

Article Section and Item	Item No.	Recommendation(s)
Title and abstract	1	Explicitly state that this is a "descriptive study" in the title or the abstract.
	2	Summarize the target population and provide an informative and balanced summary of estimated disease occurrence in the abstract.
Introduction		
Background/rationale	3	State the motivation for the study, including, where relevant, the action that might be informed by the results.
Objectives	4	State the descriptive estimand, explicitly including: (a) the target population (who would be affected by any decisions made as a result of the study?); (b) the health state to be summarized; (c) the measure of occurrence; and (d) any stratification variables, if applicable.
Methods		
Study design	5	(a) State whether the study is cross-sectional or longitudinal. (b) Restate the measure of occurrence being targeted. (c) If the study is longitudinal, specify the time origin and follow-up period for the measure of occurrence; if the study is cross-sectional, specify the time anchor at which the health state is summarized for individuals.
Setting	6	Describe any relevant features of the place and time in which the target population resides and across which data were collected.
Participants	7	(a) Describe the target population thoroughly in terms of person, place, and time. (b) Describe sampling into the study population (whether sampling was explicit or implicit, e.g., by inclusion in an administrative database); this includes eligibility criteria (see recommendations on data sources in item 10 below). (c) Describe any restrictions on the analytical sample.
Outcome(s)	8	(a) State when and how the outcome is measured. (b) Include estimates or discussion of the sensitivity and specificity of the study outcome definition relative to the gold standard. (c) List secondary outcomes or competing events of interest.
Covariates	9	Specify any stratification or adjustment variables—clearly define how variables were collected or constructed.
Data sources/measurement	10	Clearly delineate any inclusion/exclusion criteria for membership in the data source, including the original purpose for which the data were collected, if not for the study at hand.
Bias	11	Describe any assumptions or methods used to extrapolate data from the analytical sample to the study population and from the study population to the target population.
Statistical methods	12	(a) Describe the primary statistical methods used to estimate the measure of disease occurrence being targeted; discuss assumptions of that method in light of data limitations (e.g., assumption of independent censoring for people lost to follow-up). (b) If any adjustment/standardization will be done, state the goal of such adjustment.
Results		
Participants	13	Report numbers of individuals at each study stage (this is likely to be approximate for the target population); consider summarizing this information in a flow diagram.
Descriptive data	14	(a) Report on the characteristics of the analytical sample in a "Table 1." (b) Indicate the number of participants with missing data for each variable used in the analysis. (c) If any weighting or imputation is done to reconstruct the study sample or target populations, include columns for those populations.
Outcome data	15	(a) Present an overall (unstratified) estimate of the measure of occurrence of interest. (b) Report "crude" (raw data in the analytical sample) and (if applicable) "corrected" (after any weighting or imputation) estimates.
Other analyses	16	Present prespecified stratum-specific or adjusted/standardized results.

Table continues

Table 1. Continued

Article Section and Item	Item No.	Recommendation(s)
Discussion		
Key results	17	Summarize key results with reference to the study objectives.
Limitations	18	Summarize potential sources of selection bias and measurement error and any attempts to mitigate these biases. Discuss both the direction and magnitude of any potential bias. Integrating quantitative bias analysis into the study to guide these discussions is encouraged.
Interpretation	19	(a) Avoid causal interpretations of descriptive results; avoid overinterpreting stratum-specific differences in measures of occurrence. (b) Describe how results of this study might inform or improve public health or clinical practice.

mean value, etc.). While much discussion focuses on the most common scenarios (e.g., dichotomous outcomes), this framework is intended to be applied to descriptive analyses for any combination of study designs, outcomes, and estimands.

A WELL-DEFINED QUESTION

We start with the premise that good epidemiologic questions are impactful and well-defined. An impactful question, if answered, would lead to knowledge that could inform action in the population it concerns (7). A well-defined question should be stated with enough specificity and clarity that answering it is at least theoretically possible.

A well-defined research question (causal or descriptive) states: 1) the target population, characterized by person and place, and anchored in time; 2) the outcome, event, or health state or characteristic; and 3) the measure of occurrence that will be used to summarize the outcome (e.g., incidence, prevalence, average time to event, etc.). A causal question requires specifying additional components, such as exposures and covariates that are thought to be confounders, effect modifiers, or mediators. For descriptive questions, consideration of additional variables is optional, but if auxiliary variables will be considered, a well-defined descriptive question will 4) prespecify any other variables of interest and how they will be considered (e.g., to characterize the population, as a stratification factor to characterize the outcome distribution, or as a "nuisance" variable that we would like to adjust for or standardize over). For a descriptive question, indiscriminate adjustment for these other variables can lead to uninterpretable results that may mislead (9); as such, researchers should be clear as to the purpose of adjustment in descriptive studies, understand the implications of such adjustments, and be cautious in interpreting adjusted statistics (10).

Example: We illustrate application of this framework to description of one portion of the human immunodeficiency virus (HIV) care continuum (11): What was the prevalence of viral suppression on December 31, 2019, among adults living with HIV who had been linked to HIV care (i.e., saw a clinician who was aware of their HIV status and had the ability to prescribe antiretroviral therapy) in the United States? We will explore specific components of this question to make it more well-defined (and tie those components to analytical decisions) below.

SPECIFYING THE TARGET POPULATION (AND ITS RELATIONSHIP TO THE STUDY SAMPLE)

For a descriptive question, we define the target population as the group in which we would like to characterize the distribution of the outcome. The choice of target population is directly linked to the purpose of asking the question. The target population might be, for example, the population for which we will be providing public health services. The target population is not necessarily enumerated (in contrast to a cohort or a sample), but we do need to be able to define membership in terms of person, place, and time (here, time is used to define membership in the target population and does not relate directly to measurement of the outcome). For our example question, the target population is everyone living in the United States (*place*) who was aged $\geq$18 years, was infected and diagnosed with HIV, and attended $\geq$1 clinical visit for HIV care with a clinician who was aware of their infection and could prescribe antiretroviral medication (*person*) before December 31, 2019, and was alive through December 31, 2019 (*time*).

A well-defined question specifies the target population a priori. When data are available on a full census of the target population (e.g., through administrative records or public health surveillance), no sampling is needed. However, when data on the entire population cannot be obtained, we rely on data from a sample of the target population or a population that we hope is sufficiently representative of the target population with respect to both measured and unmeasured characteristics. The study sample is the enumerated set of individuals whose information is captured in a data set, among whom we attempt to measure occurrence of the outcome (after inclusion and exclusion criteria have been applied, if data were not collected using these criteria (e.g., administrative data)). Many descriptive and causal questions

are answered using convenience samples without a clear sampling frame (e.g., people recruited using Web-based surveys, frequent clinic attendees, or people who sought medical care in a particular hospital system) and implicitly assume that the study sample is a random sample (perhaps conditional on covariates with known sampling probabilities) of the target population. Achieving a representative sample may involve considerable work and may be very resource-intensive (12). However, use of convenience samples often results in study samples that are different from the target population in unmeasurable ways, particularly when subjects must actively seek out or opt into participation (13).

On the topic of sampling and selection, it is also useful to define the analytical sample as a proper subset of the study sample in which disease occurrence is measured given practical limitations (e.g., excluding individuals in the study sample who are missing information on the outcome). We might use information from the analytical sample to attempt to quantify disease occurrence in the study sample, but we must rely on assumptions to do so (e.g., assuming data are missing at random and imputing missing data or reweighting study participants with complete data). For valid inferences, the incidence of the outcome in the sample must be able to stand in for the incidence in the target population. Here, the "sample" is either the analytical sample or the study sample represented by the analytical sample after any attempts to handle missing data. Given the many practical challenges enumerated above, the samples we rely on in our studies are rarely representative of the target population. If the distribution of risk factors for the health state differs between the study sample and the target population, we have a lack of generalizability (14–16); the absolute value (risk, prevalence, rate) of the outcome in the sample will differ from what we would have observed in the target population. Without applying quantitative approaches to generalize data from the sample to the target population, descriptive results will be biased. Except in special cases (e.g., when the selected estimand is the one scale on which effect measure modification is absent), if absolute measures differ between the sample and the target, most contrasts of the outcome across exposure groups in the sample will also be biased for the same contrasts in the target population (causal results will be biased) (14–16). If the underlying joint distribution of all causes of the outcome differs between the analytical sample and the study sample, we have selection bias (17, 18). To recover an estimand relevant to the target population from an analytical sample with a different distribution of causes of the outcome, stratification and standardization methods may be appropriate.

Example: Recall that the target population is everyone living in the United States who had been linked to clinical care for HIV before December 31, 2019. There is mandated reporting in the United States of new HIV diagnoses and HIV viral load test results to public health surveillance agencies under national notifiable disease regulations, and the Centers for Disease Control and Prevention aggregates these data from all states and dependent areas. This might seem like a census of the target population. However, despite these mandates, not all diagnoses are reported, and people who move across state lines may be double-counted because of challenges with deduplication. Thus, the number of people with HIV infection may be inaccurate. Additionally, data rely on HIV viral load and CD4 cell-count laboratory tests as a proxy for clinical visits, and the proxy is imperfect (19, 20); thus, we cannot accurately apply the second inclusion criterion for target population membership: linkage to clinical care. Alternatively, we might use data from the North American AIDS Cohort Collaboration on Research and Design (NA-ACCORD) (21) or another clinical cohort of people with HIV who have been linked to care. However, clinical cohort studies are often nested within academic medical centers, where the quality of care and wraparound services may differ (and thus the probability of the outcome, viral suppression, may differ), and may have stricter enrollment criteria (to preserve study resources) than we have used to define linkage to care for our target population.

There are other options for study samples we might try to leverage. We might even choose to estimate the parameter of interest in multiple samples and triangulate the results. The point is that there is rarely a single, perfect, existing study sample that can stand in for the target population. Therefore, if we wish to use existing data, identifying ways in which the study sample and the target population differ provides a framework for thinking about sources of bias and how we might adjust the estimate for better inferences.

INTERMISSION: MISSING DATA

A theme of many threats to descriptive and causal epidemiologic inference is that they can often be cast as missing-data problems (22). The ideal data set for answering our descriptive epidemiologic question includes a row for everyone in the target population and columns with values for the outcome and any covariates of interest. When the study sample is not a census of the target population, anyone in the target population who is not in the study sample will have missing data in some, if not all, columns. Indeed, without a clear sampling frame, we do not even know how many rows are missing from our ideal data set (and we cannot quantify the amount of missing data from this ideal study). Analyzing the study sample as if it were a random sample of the target population is akin to assuming that data are missing completely at random. If, instead, it is plausible to assume that data are missing at random conditional on covariates that are available for target population members who were not selected for the study sample, we could reweight or standardize the study sample to represent the full target population.

Example: The surveillance data include everyone in the target population (age $\geq$18 years, alive, diagnosed with HIV, and $\geq$1 HIV care visit before December 31, 2019), but they also include some people who are not in the target population (they include people who did not make $\geq$1 HIV care visit with a clinician who might prescribe antiretroviral medications), and we are unable to definitively identify people in the surveillance data who do not meet the inclusion criteria for the target population (we have to rely on laboratory tests as a proxy for clinical visits) (19). However, the surveillance data likely are closer to representing the target population than the

NA-ACCORD data (which do not include everyone in the target population, although they do not include anyone who should be excluded from the target population). Therefore, we might use surveillance data for our primary analyses, but we might conduct secondary analyses that leverage the relative strengths of the different study samples and, for example, reweight NA-ACCORD data that include visits to resemble the target population implied by the surveillance data.

DEFINING THE OUTCOME

To describe the occurrence, frequency, or relative frequency of an outcome, we need an unambiguous definition of that outcome, and we must be able to apply that definition in our data. In the absence of a gold standard or the ability to apply that gold standard due to data or resource constraints, we must understand how imperfect sensitivity and specificity might affect our results. Measurement error has previously been described as a missing-data problem (22) in which the true outcome is missing and we overwrite that missing value with a mismeasured outcome. To the extent to which the mismeasured outcome is a poor substitute for the true outcome, our inferences will be biased.

Example: Our outcome is "viral suppression" on December 31, 2019, but there is no single, standard threshold for suppression. Prior studies have used plasma HIV RNA levels of <20, <50, <200, or <400 copies/mL (23). Lower thresholds will result in a lower estimate of the prevalence of viral suppression; for example, in an HIV clinical cohort in Baltimore, Maryland, the proportion of patients estimated to have a suppressed viral load in a given year from 2010 to 2018 was 75% if the threshold for suppression was set at <20 copies/mL but 89% if the threshold was set at <400 copies/mL (24). Failure to suppress viral load below a lower threshold may also be a more sensitive indicator of subsequent morbidity and mortality (24–28), but suppression below a higher threshold is more relevant as an indicator of an individual's transmission potential (29, 30), so our choice of threshold may depend on how our results will be used. Additionally, not everyone in either of our candidate study samples will have had a viral load measurement on December 31, 2019, exactly. Typically, researchers accept viral loads measured within a time window around some key date as indicative of the viral load on that key date. We must decide how wide a window we are willing to use to answer our question. The width we are willing to tolerate might depend on how frequently we anticipate viral load changes in the population. A wider window risks assigning a viral load value to December 31 that is inaccurate because viral load has changed since measurement, while a narrower window will result in a larger proportion of the cohort with a missing viral-load value.

SPECIFYING A MEASURE OF OCCURRENCE

We have multiple options for measures of occurrence, and like the proverbial blind men feeling the elephant, our choice of measure of occurrence might give us only part of the complete picture about the distribution of the outcome in the target population. Incidence tells us something about how frequently an event occurs over time. There are multiple measures of incidence; in the interest of space, we will restrict our discussion to risks and rates. If individuals are not followed over time and the event can recur, it may be difficult to distinguish the number of affected individuals from the number of events. Prevalent outcomes are often not of interest in causal investigations, as temporality is more challenging to determine and reverse causation is a potential problem. In addition, survival bias might affect results when considering prevalent exposures (31, 32). Finally, prevalence is a function of the incidence of the condition and its duration, such that, if incidence is what is relevant to the question at hand, prevalence might be a misleading proxy. However, for descriptive questions designed to inform public-health planning for secondary or tertiary prevention measures, prevalence might be the most relevant measure of occurrence, as it reflects the population of people who might access those services.

Risk (the proportion of people free from disease at baseline who develop the outcome during the study period) is the foundation of many causal epidemiologic studies (33), particularly as the target trial framework (1) has gained in popularity. Risk is arguably the most easily interpretable measure of disease occurrence for the general public (33). We discuss rates (the number of events divided by a sum of person-time) as an alternative measure of incidence in a few paragraphs. Two complications for obtaining valid estimates of either measure of incidence, however, are competing events and incompletely observed person-time (left-truncation and right-censoring).

Competing events are events that preclude the event of interest from occurring and are theoretical if not practical problems for all outcomes other than all-cause mortality (34). In the presence of competing events, we have the option to report the conditional or unconditional risk (i.e., cumulative incidence function) (35). The conditional risk is the proportion of people free from disease at baseline that we would expect to develop the outcome during the study period if all competing events were prevented without changing the hazard of the event of interest; it is the risk "conditional" on removal of the competing event. It is estimated by censoring persons who experience a competing event and is the first and sometimes only estimand of risk that students of epidemiology are taught (36). It is also implied by the exponential formula for converting rates to risks. However, complete removal of the competing event is a hypothetical intervention, and the conditional risk is the risk under that often-infeasible intervention. If our goal is to describe the world as it exists, absent hypothetical interventions, the cumulative incidence function is recommended when the number of competing events is nontrivial (37). The cumulative incidence function (or, as is implied but is a less commonly used term, the unconditional risk) is the proportion of people free from disease at baseline who would develop the outcome of interest during the study period in the real world in which a competing event might remove them from follow-up and preclude them from ever developing the outcome of interest.

Risks can be calculated in the presence of late entries (left-truncation) and loss to follow-up (right-censoring) under strong assumptions about independence between entering/leaving the study and risk of the outcome (38, 39). Left-truncation and right-censoring impute outcomes for people who did not survive to enroll in the study sample and for people who are censored (38). We can adjust for possible associations between censoring and the outcome (and resultant selection bias) using inverse probability of censoring weights (40). However, the resultant risks are interpretable as the risk that would have been observed if no one were lost to follow-up (a hypothetical intervention), and will be different from the natural course if loss to follow-up was associated with the outcome in ways not captured by covariates in the weight model or if loss to follow-up itself directly altered the risk of the outcome (18, 40).

Finally, rates may occasionally be a useful measure of incidence as an alternative to risks, especially for descriptive studies. Risks are only defined relevant to a population free of, and biologically at risk for, the outcome at a particular time origin. When we would like to describe incidence across a time metric along which not all people were biologically at risk at the time origin, rates can appropriately exclude person-time not at risk and allow for reporting of smoothed incidence estimates. For example, when describing temporal trends for the incidence of HIV diagnoses since the beginning of the epidemic in the 1980s, there will be people who were not born (not at risk for the outcome) in the 1980s who should be counted in the target population in the 2010s. Perhaps in an idealized descriptive study, we would report the daily risk of HIV diagnosis restricted to people who were alive and at risk for HIV diagnosis at the start of each day. However, across 3 decades this may be computationally intensive and impractical given the granularity of data collection and reporting. We might instead report weekly, monthly, or yearly HIV diagnosis risk, but the wider the time interval across which we measure risk becomes, the greater the number of people in our target population who are not at risk at the start of the interval. How should we treat people born in December 1990 when calculating the risk of HIV diagnosis in 1990? In contrast, if we are willing to assume that the rate of HIV diagnosis across a calendar year is approximately constant, or if we assume that the average rate is a reasonable representation of the incidence in that year, rates could appropriately exclude person-time in which people are not biologically at risk. The assumption of a constant rate or the acceptability of an average rate for answering the study question should be plausible across the time intervals chosen, or time should be further discretized. Another benefit of rates is that they are straightforward to estimate when we do not have individual-level data, which is more common in descriptive analyses than in causal or predictive epidemiologic analyses. For example, rates are the standard measure of incidence used for notifiable diseases, where health departments count case reports to get the numerator and use midyear census estimates for the denominator.

Example: We have clearly specified in our research question that we are interested in the prevalence of viral suppression on December 31, 2019. People in our study sample with no viral load measurement in 2019 are lost to follow-up. Viral suppression is influenced by access to health care and is only possible if people are receiving antiretroviral therapy (except, in rare cases, for elite controllers) (41). In this setting, people who are lost to follow-up may have transferred to another clinic and may still be receiving treatment (if we are using NA-ACCORD data) or may have moved out of the jurisdiction (if we are using surveillance data), and we might assume that they have the same probability of viral suppression as people with a viral load measurement (censoring is appropriate; equivalently, we can restrict analyses to people with a measured viral load) (24). Alternatively, people who do not have a viral load measurement may have dropped out of clinical care and may not have access to antiretroviral therapy. The probability of viral suppression among these individuals is near 0 (we might think of loss to follow-up as a competing event and assign a value of "not suppressed" to persons who are lost to follow-up) (42). Understanding the assumptions and implications of different analytical decisions for these people is critical for making the right inference about the prevalence of the outcome.

THE ROLE OF COVARIATES

When describing the prevalence or incidence of an outcome, we sometimes want to characterize the people who got the outcome according to covariates. Alternatively, we may want to account for nuisance variables, such as factors that differ between the study sample and the target population or between groups we plan to stratify by. When characterizing groups with the highest incidence of the outcome, bivariate results can make it challenging to understand how covariates interact to determine the distribution of disease. For example, if the prevalence of viral suppression is lower for cisgender women than for cisgender men and lower for Black patients than for White patients (43), what would we expect to see regarding the prevalence of viral suppression for cisgender White women relative to cisgender Black men? Stratifying on multiple variables simultaneously might be helpful in this setting, or we may want to employ theoretical models (e.g., conceptual frameworks for how variables influence risk of the outcome) or statistical strategies (e.g., supervised machine learning) to identify the most important variables if there are not enough data to stratify on all variables of interest. Conversely, when trying to understand whether one covariate is associated with the distribution of disease independently or merely because of its correlation with another covariate, a common approach is to put all covariates into a single model. However, this approach can lead to incorrect interpretations of the results and inappropriate recommendations for actions (44). Adjustment implies an intervention on the data and a distortion of reality—for example, "Would Black people still have lower prevalence of viral suppression if they had the same distribution of HIV acquisition risk factors as White people?". Inappropriate adjustment may understate the magnitude of disparities (45) and adjusted statistics are prone to be interpreted causally, which could lead to inappropriate recommendations (9). We endorse reporting and primary interpretation of unadjusted

results for descriptive studies and clear justification and proper interpretation in cases where adjustments are made.

CONCLUSIONS

Descriptive epidemiologic studies seek to characterize what is happening in the world to inform public health priorities, target interventions, and occasionally contrast with counterfactual scenarios to estimate intervention effects (46, 47). Descriptive studies have value in their own right and not merely as stepping stools toward causal inference. Characterizing what is happening in the world requires that we be very clear about the particular slice of the world and the specific outcome we hope to study. Generalizability and selection biases can bias descriptive studies when study participation is associated with the outcome. Measurement error can bias descriptive studies when we do not use, or there is no gold-standard measure of, the outcome. Different measures of occurrence will provide different pictures of what is happening in the world. Censoring people who have a competing event or adjusting for covariates implies interventions on the data such that the results are a distorted version of reality. These are all basic epidemiologic principles that also affect the success of our attempts at causal effect estimation. Performing rigorous descriptive studies that accurately estimate a parameter of interest and are interpretable to clinicians and policy-makers will improve public health.

ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland, United States (Catherine R. Lesko); Departments of Epidemiology and Global Health, School of Public Health, Boston University, Boston, Massachusetts, United States (Matthew P. Fox); and Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States (Jessie K. Edwards).

This work was supported by grants K01 AA028193, K01 AI125087, and R01 AI157758 from the National Institutes of Health.

Conflict of interest: none declared.

REFERENCES

1. 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.
2. Petersen ML, van der Laan MJ. Causal models and learning from data: integrating causal modeling and statistical estimation. *Epidemiology.* 2014;25(3):418–426.
3. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. *Int J Surg.* 2014;12(12):1495–1499.
4. Robins JM. Data, design, and background knowledge in etiologic inference. *Epidemiology.* 2001;12(3):313–320.
5. 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.
6. Petersen ML. Commentary: applying a causal road map in settings with time-dependent confounding. *Epidemiology.* 2014;25(6):898–901.
7. Fox MP, Edwards JK, Platt R, et al. The critical importance of asking good questions: the role of epidemiology doctoral training programs. *Am J Epidemiol.* 2020;189(4):261–264.
8. Holland PW. Statistics and causal inference. *J Am Stat Assoc.* 1986;81(396):945–960.
9. Tennant PWG, Murray EJ. The quest for timely insights into COVID-19 should not come at the cost of scientific rigor. *Epidemiology.* 2021;32(1):e2.
10. Kaufman JS. Statistics, adjusted statistics, and maladjusted statistics. *Am J Law Med.* 2017;43(2-3):193–208.
11. Gardner EM, McLees MP, Steiner JF, et al. The spectrum of engagement in HIV care and its relevance to test-and-treat strategies for prevention of HIV infection. *Clin Infect Dis.* 2011;52(6):793–800.
12. Lee KK, Fitts MS, Conigrave JH, et al. Recruiting a representative sample of urban South Australian Aboriginal adults for a survey on alcohol consumption. *BMC Med Res Methodol.* 2020;20(1):183.
13. Offord C. How (not) to do an antibody survey for SARS-CoV-2. *Scientist.* https://www.the-scientist.com/news-opinion/how-not-to-do-an-antibody-survey-for-sars-cov-2-67488. Published April 28, 2020. Accessed April 8, 2022.
14. Lesko CR, Buchanan AL, Westreich D, et al. Generalizing study results: a potential outcomes perspective. *Epidemiology.* 2017;28(4):553–561.
15. Dahabreh IJ, Robertson SE, Tchetgen EJ, et al. Generalizing causal inferences from individuals in randomized trials to all trial-eligible individuals. *Biometrics.* 2019;75(2):685–694.
16. Cole SR, Stuart EA. Generalizing evidence from randomized clinical trials to target populations: the ACTG 320 Trial. *Am J Epidemiol.* 2010;172(1):107–115.
17. Westreich D. Berkson's bias, selection bias, and missing data. *Epidemiology.* 2012;23(1):159–164.
18. Hernán MA. Invited commentary: selection bias without colliders. *Am J Epidemiol.* 2017;185(11):1048–1050.
19. Rebeiro PF, Althoff KN, Lau B, et al. Laboratory measures as proxies for primary care encounters: implications for quantifying clinical retention among HIV-infected adults in North America. *Am J Epidemiol.* 2015;182(11):952–960.
20. Lesko CR, Sampson LA, Miller WC, et al. Measuring the HIV care continuum using public health surveillance data in the United States. *J Acquir Immune Defic Syndr.* 2015;70(5):489–494.
21. Gange SJ, Kitahata MM, Saag MS, et al. Cohort profile: the North American AIDS Cohort Collaboration on Research and Design (NA-ACCORD). *Int J Epidemiol.* 2007;36(2):294–301.
22. Edwards JK, Cole SR, Westreich D. All your data are always missing: incorporating bias due to measurement error into the potential outcomes framework. *Int J Epidemiol.* 2015;44(4):1452–1459.
23. McMahon JH, Elliott JH, Bertagnolio S, et al. Viral suppression after 12 months of antiretroviral therapy in

low- and middle-income countries: a systematic review. *Bull World Health Organ.* 2013;91(5):377–385E.
24. Lesko CR, Chander G, Moore RD, et al. Variation in estimated viral suppression associated with the definition of viral suppression used. *AIDS.* 2020;34(10):1519–1526.
25. Hermans LE, Moorhouse M, Carmona S, et al. Effect of HIV-1 low-level viraemia during antiretroviral therapy on treatment outcomes in WHO-guided South African treatment programmes: a multicentre cohort study. *Lancet Infect Dis.* 2018;18(2):188–197.
26. Elvstam O, Medstrand P, Yilmaz A, et al. Virological failure and all-cause mortality in HIV-positive adults with low-level viremia during antiretroviral treatment. *PLoS One.* 2017; 12(7):e0180761.
27. Antiretroviral Therapy Cohort Collaboration, Vandenhende MA, Ingle S, et al. Impact of low-level viremia on clinical and virological outcomes in treated HIV-1-infected patients. *AIDS.* 2015;29(3):373–383.
28. Laprise C, de Pokomandy A, Baril J-G, et al. Virologic failure following persistent low-level viremia in a cohort of HIV-positive patients: results from 12 years of observation. *Clin Infect Dis.* 2013;57(10):1489–1496.
29. Lesko CR, Lau B, Chander G, et al. Time spent with HIV viral load >1500 copies/mL among persons engaged in continuity HIV care in an urban clinic in the United States, 2010–2015. *AIDS Behav*. 2018;22(11):3443–3450.
30. Quinn TC, Wawer MJ, Sewankambo N, et al. Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai Project Study Group. *N Engl J Med.* 2000; 342(13):921–929.
31. Prentice RL, Chlebowski RT, Stefanick ML, et al. Estrogen plus progestin therapy and breast cancer in recently postmenopausal women. *Am J Epidemiol.* 2008;167(10): 1207–1216.
32. Lund JL, Richardson DB, Stürmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. *Curr Epidemiol Rep.* 2015;2(4):221–228.
33. Cole SR, Hudgens MG, Brookhart MA, et al. Risk. *Am J Epidemiol.* 2015;181(4):246–250.
34. Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. *Am J Epidemiol.* 2009;170(2): 244–256.
35. Edwards JK, Hester LL, Gokhale M, et al. Methodologic issues when estimating risks in pharmacoepidemiology. *Curr Epidemiol Rep.* 2016;3(4):285–296.
36. Rothman KJ, Lash TL, VanderWeele TJ, et al. Measures of occurrence. In: *Modern Epidemiology*. 4th ed. Philadelphia, PA: Wolters Kluwer N.V.; 2021:53–77.
37. Cole SR, Lau B, Eron JJ, et al. Estimation of the standardized risk difference and ratio in a competing risks framework: application to injection drug use and progression to AIDS after initiation of antiretroviral therapy. *Am J Epidemiol.* 2015;181(4):238–245.
38. Cole SR, Edwards JK, Naimi AI, et al. Hidden imputations and the Kaplan-Meier estimator. *Am J Epidemiol.* 2020; 189(11):1408–1411.
39. Lesko CR, Edwards JK, Cole SR, et al. When to censor? *Am J Epidemiol.* 2018;187(3):623–632.
40. Howe CJ, Cole SR, Lau B, et al. Selection bias due to loss to follow up in cohort studies. *Epidemiology.* 2016;27(1): 91–97.
41. Okulicz JF, Marconi VC, Landrum ML, et al. Clinical outcomes of elite controllers, viremic controllers, and long-term nonprogressors in the US Department of Defense HIV Natural History Study. *J Infect Dis.* 2009;200(11): 1714–1723.
42. Edwards JK, Lesko CR, Herce ME, et al. Gone but not lost: implications for estimating HIV care outcomes when loss to clinic is not loss to care. *Epidemiology.* 2020;31(4): 570–577.
43. Centers for Disease Control and Prevention. *Monitoring Selected National HIV Prevention and Care Objectives by Using HIV Surveillance Data—United States and 6 Dependent Areas, 2019*. (*HIV Surveillance Supplemental Report*, vol. 26, no. 2). Atlanta, GA: Centers for Disease Control and Prevention; 2021. https://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-report-vol-26-no-2.pdf. Accessed November 29, 2021.
44. Westreich D, Greenland S. The table 2 fallacy: presenting and interpreting confounder and modifier coefficients. *Am J Epidemiol.* 2013;177(4):292–298.
45. Zalla LC, Martin CL, Edwards JK, et al. A geography of risk: structural racism and COVID-19 mortality in the United States. *Am J Epidemiol.* 2021;190(8):1439–1446.
46. Westreich D. From exposures to population interventions: pregnancy and response to HIV therapy. *Am J Epidemiol.* 2014;179(7):797–806.
47. Edwards JK, Cole SR, Lesko CR, et al. An illustration of inverse probability weighting to estimate policy-relevant causal effects. *Am J Epidemiol.* 2016;184(4): 336–344.