

# The Critical Importance of Asking Good Questions: The Role of Epidemiology Doctoral Training Programs

Matthew P. Fox\*, Jessie K. Edwards, Robert Platt, and Laura B. Balzer

\* Correspondence to Matthew P. Fox, 801 Massachusetts Avenue, Boston, MA 02218 (e-mail: mfox@bu.edu).

*Initially submitted April 30, 2019; accepted for publication September 30, 2019.*

Epidemiologic methods have advanced tremendously in the last several decades. As important as they are, even the most sophisticated approaches are unable to provide meaningful answers when the user lacks a clear study question. Yet, instructors have more and more resources on how to conduct studies and analyze data but few resources on how to ask clearly defined study questions that will guide those methods. Training programs have limited time for coursework, and if novel statistical estimation methods become the focus of instruction, programs that go this route may end up underemphasizing the process of asking good study questions, designing robust studies, considering potential biases in the collected data, and appropriately interpreting the results of the analysis. Given the demands for space in curricula, now is an appropriate time to reevaluate what we teach epidemiology doctoral students. We advocate that programs place a renewed focus on asking good study questions and following a comprehensive approach to study design and data analysis in which questions guide the choice of appropriate methods, helping us avoid methods for methods' sake and highlighting when application of a new method can provide the opportunity to answer questions that were intractable with traditional approaches.

causal inference; epidemiologic methods; novel methods; study questions; teaching; training

To gain important and actionable insights into the health of populations, epidemiologists must start with clearly defined study questions. There is no doubt that important insights have been identified through simple observation outside the formal scientific process; however, rigorous evaluations require clearly defined questions that can be answered and rigorous hypotheses that can be tested. Such questions are best when they are of importance to advancing knowledge that will improve the human condition (i.e., they are questions of importance), but they must also be specified in such a way that we know what the results are telling us about the world we inhabit. For this to work, the methods we use must follow logically from the question we ask and not the other way around. And yet, at least in some programs, instructors have focused more and more resources on how to conduct studies and analyze data but few resources on how to ask clearly defined study questions that will guide those methods. Given the demands for space in curricula, now is an appropriate time to reevaluate what we teach epidemiology doctoral students (1–4). Specifically, we advocate that programs place a renewed focus on asking good study questions.

## THE CRITICAL IMPORTANCE OF ASKING GOOD STUDY QUESTIONS

Good study questions come from a clear statement of the problem at hand and tend to focus on problems that can lead to knowledge of relevance to the population. For the same exposure outcome pair, epidemiologists tend to ask 3 main types of questions: descriptive, predictive, and causal (5). All are critical to solving public health problems, all should be valued within epidemiologic training programs, and students should be taught how to ask each and the differences between them. For example, suppose we are interested in a treatment for individuals diagnosed with a particular condition. We might want to ask whether persons who were exposed had survival rates different from those in persons who were not. We might also want to ask whether we can predict the probability of survival as a function of baseline factors including the exposure and other covariates. Finally, we might ask, would patients have longer survival, on average, if they received the exposure compared with not receiving it? Note that only the last question can be thought of as a causal question, seeking to

answer what the outcome would have been if the conditions had changed and, therefore, provides guidance on clinical or policy decisions. The first question might be confused for a causal question, but it is a descriptive question because it simply asks to compare outcomes for those who received treatment and those who did not, rather than aiming to learn what would have happened if all patients had received treatment or not. In the first case, we would not be concerned about controlling confounding, among other potential biases. In other words, the descriptive question only compares what actually happened to different groups of people, whereas the causal question seeks to understand what would have happened to the same people under different interventions.

In all cases, explicit training on asking good study questions becomes critical. For causal research questions, specifying a target trial (6–10) or using the causal roadmap (11–15) are helpful approaches to ensure study design and analytic approaches follow appropriately from those well-specified questions (16, 17) and can help ensure the scientific question is driving the analysis. For descriptive and predictive questions, an analogous roadmap could also be developed to guide the development of the study and the resulting analyses and interpretation. Whatever approach is used, the methods should follow from the question and not the other way around.

## **WHAT TO EMPHASIZE IN DOCTORAL TRAINING PROGRAMS**

Asking good study questions and choosing appropriate statistical estimation methods to answer these questions are difficult endeavors and require appropriate training. Early training tends to emphasize regression modeling and may leave students with the impression that statistical methods come before basic principles of study design or that answering questions is about deciding which method to use for a particular type of data. In addition, we, as trainers, often assume students entering doctoral programs already have sufficient training in how to ask a study question, especially if they come to programs with some work experience; yet, this may not always be the case. Explicit training is needed to translate an interesting idea that arises in a work environment to a study question.

Novel statistical estimation methods provide the promise of estimating causal effects in ways that we never could before, getting better prediction and improving descriptive epidemiology, and this is useful for improving public health. In some cases, failure to use novel methods will give us biased results. In some cases, novel methods allow us to answer more useful or policy-relevant questions. Therefore, it is essential that we teach these tools to our trainees alongside study design and implementation. At the same time, novel statistical estimation methods must help ensure our scientific question drives the analysis and not the other way around. Furthermore, we need to ensure that use of causal methods does not drive overinterpretation of study results, where the presence of strong, uncontrolled biases leads us to draw conclusions that are not justified from the data, even after application of causal methods.

## **TIME FOR A RETHINK OF TRAINING PROGRAMS?**

What are potential solutions to these competing demands on training time? First, we propose that following a comprehensive approach to study design and data analysis is essential. For causal questions, this may be accomplished by following a causal roadmap approach (11, 12), which advocates for ensuring etiologic study questions are translated into well-defined causal effects; our knowledge of the study design is expressed into a causal model; the assumptions needed to answer that causal question with the observed data are explicitly stated and transparently assessed; the resulting statistical analysis implemented is the best choice to answer that underlying scientific question while avoiding introducing new assumptions; and the resulting point estimates are correctly interpreted (and not overinterpreted). Or it may involve focusing students on using target trials (6–10), in which, when possible, students conducting observational studies first design the hypothetical randomized trial they would do if they could, then design their observational studies to mimic the hypothetical trial as closely as possible. Such approaches are logical ways of defining a clear study question and allowing the methods to follow from the question proposed. These approaches also clarify how limitations in the available data may require us to refine our scientific question, collect new data, and temper our interpretations. It also makes clear that multiple questions can be asked of the same data set, and forces the investigators to be clear about what question they wish to answer. For description and prediction questions, a roadmap approach may be just as valuable because it forces us to specify a clear question, identify how the parameters will be used, identify a clear target population, guard against sources of bias like selection bias, and identify an appropriate estimand and estimation method.

It is important to note that such a rethink may or may not come at the expense of teaching some novel statistical estimation methods. In fact, not all programs have this issue of overemphasizing novel statistical estimation methods, and such programs may need to increase their training in methods. At the same time, we would encourage such programs to revisit how they currently teach asking good study questions and ensure that deficits are addressed and any changes to curriculum to emphasize novel methods do not come at the expense of training in asking good study questions. Some programs that go through this process may identify that once students are oriented to a questions-first approach, they require more training in such methods to be able to answer the relevant questions. Thus, such a reevaluation of doctoral training programs could logically lead to an increase in methods training. In contrast, some programs that go through this process may realize that carving out time to train students to ask good questions requires sacrificing some training in some novel statistical estimation techniques. Furthermore, we note that this tension is not a new problem or one that will go away in the future. Programs should be in a state of constant refinement in this regard, aiming to maintain balance while managing ever-increasing methods development.

Second, we encourage the use of case-based teaching, where a study design and data analysis approach are discussed

as a whole and are not left to different courses or taught in isolation. This would encourage a focus on thinking about how most existing and novel study designs are simply ways to approximate the idealized study. When training in statistical estimation methods is divorced from training in study design, students may be led to believe that analysis takes priority and design issues are less important. Although rigorous training in statistical estimation techniques is critical, application of these statistical techniques should occur within the context of well-defined study questions in doctoral training programs.

Third, we encourage openness to the idea of teaching both study design and analysis (i.e., how to ask a study question and how to answer it) by first placing a strong emphasis on simple stratification and standardization approaches and using the intuition gained in learning these approaches to guide implementation of regression and machine-learning approaches. This would have the advantage of emphasizing that questions and methods cannot be separated and regression (or more flexible semiparametric/nonparametric approaches) cannot begin without a clearly defined question.

Finally, we recommend focusing on toolboxes rather than tools—that is, how to work with multiple implements and select the right one for each problem, rather than honing skill with individual tools. Learning to use a complex method on a well-defined problem in a clean data set is much more straightforward than learning first how to specify a well-defined problem and then select and implement the appropriate analyses on a real data set. An apprenticeship model may be of use here. A tradesperson learns to complete a project by first specifying the problem and then selecting the appropriate tool for each task under the supervision of an experienced colleague. We may need to work toward teaching epidemiologists and data scientists in a similar way.

## CONCLUSIONS

There is no doubt that novel statistical estimation methods are the future of epidemiology. New and emerging analytic methods enable epidemiologists to answer a host of important clinical and public health questions that were previously intractable, and to clarify the assumptions underlying the inferences generated. With a well-defined question in mind, application of new statistical estimation techniques allows epidemiologists to better emulate the idealized study by allowing them to account for selection bias (18), generalize and transport findings to relevant target populations (19), account for measurement error (20), and deal with complex confounding structures (21). We cannot and should not avoid embracing them and we should make sure our students are thoroughly versed in them. At the same time, we should be proactive in thinking through what is essential and what will help us grow as a field. Although not everyone can and should be an expert in all aspects of epidemiology, students should be able to articulate a well-defined study question and identify situations where standard statistical estimation methods fail to provide an unbiased answer. Given the limits in time to train epidemiologists, we must make clear and intentional decisions about our priorities in training the next

generation. This might mean that some well-loved statistical estimation methods (new and old ones) receive less emphasis. Whatever the approach, focusing on questions first will guide the choice of appropriate methods, helping us avoid methods for methods' sake and highlighting when application of a new method can provide the opportunity to answer questions that were intractable with traditional epidemiologic approaches.

## ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology, Boston University School of Public Health, Boston University, Boston, Massachusetts (Matthew P. Fox); Department of Global Health, Boston University School of Public Health, Boston University, Boston, Massachusetts (Matthew P. Fox); Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina (Jessie K. Edwards); Department of Pediatrics, McGill University, Montreal, Quebec, Canada (Robert Platt); Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Quebec, Canada (Robert Platt); and Department of Biostatistics and Epidemiology, School of Public Health and Health Sciences, University of Massachusetts-Amherst, Amherst, Massachusetts (Laura B. Balzer).

R.P. holds the Albert Boehringer I Chair in Pharmacoepidemiology. J.K.E. was supported by National Institutes of Health, National Institute of Allergy and Infectious Diseases grant K01AI125087.

Conflict of interest: none declared.

## REFERENCES

1. Lau B, Duggal P, Ehrhardt S. Epidemiology at a time for unity [published correction appears in *Int J Epidemiol*. 2019;48(1):321]. *Int J Epidemiol*. 2019;48(1):1366–1371.
2. Werler MM, Stuver SO, Healey MA, et al. The future of teaching epidemiology. *Am J Epidemiol*. 2019;188(5):825–829.
3. Glymour MM, Bibbins-Domingo K. The future of observational epidemiology: improving data and design to align with population health. *Am J Epidemiol*. 2019;188(5):836–839.
4. Bensyl DM, King ME, Greiner A. Applied epidemiology training needs for the modern epidemiologist. *Am J Epidemiol*. 2019;88(5):830–835.
5. Hernán MA, Hsu J, Healy B. A second chance to get causal inference right: a classification of data science tasks. *Chance*. 2019;32(1):42–49.
6. 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.
7. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available: table 1. *Am J Epidemiol*. 2016;183(8):758–764.
8. García-Albéniz X, Hsu J, Hernán MA. The value of explicitly emulating a target trial when using real world evidence: an

- application to colorectal cancer screening. *Eur J Epidemiol*. 2017;32(6):495–500.
9. Caniglia EC, Zash R, Jacobson DL, et al. Emulating a target trial of antiretroviral therapy regimens started before conception and risk of adverse birth outcomes. *AIDS*. 2018;32(1):113–120.
  10. Danaei G, García Rodríguez LA, Cantero OF, et al. Electronic medical records can be used to emulate target trials of sustained treatment strategies. *J Clin Epidemiol*. 2018; 96:12–22.
  11. Petersen ML, van der Laan MJ. Causal models and learning from data: integrating causal modeling and statistical estimation. *Epidemiology*. 2014;25(3):418–426.
  12. Balzer L, Petersen M, van der Laan M. Tutorial for causal inference. In: Buhlmann P, Drineas P, Kane M, et al., eds. *Handbook of Big Data*. Boca Raton, FL: Chapman & Hall/CRC; 2016.
  13. Tran L, Yiannoutsos CT, Musick BS, et al. Evaluating the impact of a HIV low-risk express care task-shifting program: a case study of the targeted learning roadmap. *Epidemiol Method*. 2016;5(1):69–91.
  14. Kreif N, Tran L, Grieve R, et al. Estimating the comparative effectiveness of feeding interventions in the pediatric intensive care unit: a demonstration of longitudinal targeted maximum likelihood estimation. *Am J Epidemiol*. 2017; 186(12):1370–1379.
  15. Petersen ML. Commentary: applying a causal road map in settings with time-dependent confounding. *Epidemiology*. 2014;25(6):898–901.
  16. Hernán MA. The C-word: scientific euphemisms do not improve causal inference from observational data. *Am J Public Health*. 2018;108(5):616–619.
  17. Ahern J. Start with the “C-word,” follow the roadmap for causal inference. *Am J Public Health*. 2018;108(5):621.
  18. 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.
  19. 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.
  20. Fox M, Lash T, Greenland S. A method to automate probabilistic sensitivity analyses of misclassified binary variables. *Int J Epidemiol*. 2005;34(6):1370–1376.
  21. 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;7:1393–1512.