

Studies Using Randomized Trial Data to Compare Nonrandomized Exposures

Catharine B. Stack, PhD; Anne R. Meibohm, PhD; Joshua M. Liao, MD, MSc; and Eliseo Guallar, MD, DrPH

Researchers frequently use data collected in randomized controlled trials to target questions that are beyond the scope of the original trial. For example, trial data may be used to explore the effects on outcomes of an exposure collected during the trial that was not the randomized intervention. Although such data originate from a randomized trial, the analyses do not have the protection from bias that randomization provides. Successful randomization tends to balance observed and unobserved characteristics between study groups. Comparisons of nonrandomized exposures, even when done using trial data, can be biased and require analytic approaches commonly used for observational studies (for example, multivariable adjustment, stratification, weighting, and matching).

The goal of this article is to help readers understand key questions to consider when interpreting studies that use trial data to compare nonrandomized groups (Table). This article does not focus on analyses comparing randomized groups, which may also be susceptible to biases when adherence is poor or data are incomplete (1).

THE ORIGINAL TRIAL

In the original randomized controlled trial (Helping HAND 2), 1357 hospitalized smokers, all of whom expressed a desire to quit tobacco and received in-hospital tobacco cessation counseling, were randomly assigned to receive either 3 months of standard care for tobacco cessation ($n = 677$) or an intervention with sustained care ($n = 680$) (2). Standard care included an individualized recommendation for a postdischarge smoking cessation medication plus advice to call a free telephone quitline. The intervention included an up to 90-day free supply of tobacco cessation medication, which was selected by the patient with guidance from a smoking counselor, and 5 automated, interactive telephone calls at specific intervals after discharge that encouraged adherence to the medication and offered direct transfer to a telephone quitline.

The primary outcome was 7-day abstinence from all tobacco products, including e-cigarettes, at 6 months after discharge, which was measured biochemically by testing for carbon monoxide in expired air or cotinine in mailed saliva samples. The researchers did not find a difference in abstinence between the intervention and standard care groups (16.6% vs. 15.5%; risk difference, 1.1 percentage points [95% CI, -2.8 to 5.0 percentage points]).

THE EXAMPLE: COMPARISON OF NONRANDOMIZED EXPOSURE

A post hoc analysis of these data (the e-cigarette analysis) aimed to determine whether e-cigarette use in

Table. Key Questions for Evaluating Studies Using Trial Data to Compare Nonrandomized Exposures

How does the analysis sample compare with the full, randomized trial sample? How might these differences create bias?
How does the exposure compare with the randomized intervention in the original trial? How might these differences create bias?
How does the method of analysis control for these sources of bias?
How are results from the analysis interpreted?

the 3 months after hospital discharge was associated with subsequent tobacco abstinence (3). Researchers identified 1022 patients with data about self-reported e-cigarette use in the 3 months after hospital discharge; 286 reported using e-cigarettes (exposed) and 736 did not (unexposed). As part of the analysis, the researchers created a propensity score for e-cigarette use and applied propensity score matching to form 237 pairs of patients. (For the definition of *propensity score* and other terms used in the article, see the Glossary [4].) Each pair had similar characteristics at the start of the trial, and both members of the pair were randomly assigned to the same group (sustained or standard care). One member was exposed to e-cigarette use in the 3 months after hospital discharge, and the other was not. The researchers did an analysis using these matched pairs to evaluate the association between e-cigarette use and biochemically verified 7-day tobacco abstinence at 6 months (the outcome). They found that e-cigarette users were less likely than nonusers to achieve abstinence (10.1% vs. 26.6%; risk difference, -16.5 percentage points [CI, -23.3 to -9.6 percentage points]). The Figure illustrates the design of the original trial and the analysis comparing e-cigarette users with nonusers.

HOW DOES THE ANALYSIS SAMPLE COMPARE WITH THE FULL, RANDOMIZED TRIAL SAMPLE? HOW MIGHT THESE DIFFERENCES CREATE BIAS?

The original trial enrolled 1357 patients who were randomly assigned to receive either sustained or standard care for smoking cessation. The e-cigarette analysis included 1022 patients who provided an assessment of e-cigarette use during the first 3 months after hospital discharge (75% of the original randomized sample). Patients who were lost to follow-up at 3 months were younger (mean age, 46 vs. 51 years) and were less likely to have a smoking-related disease as their primary discharge diagnosis (24% vs. 37%) than those who remained in the study, but they did not differ in the number of cigarettes smoked per day. Matching further changed the characteristics of the analysis sample. The 1:1 matching scheme identified non-e-

Glossary

- Confounder:* A variable that is related to both the exposure of interest and the outcome so that it alters (biases) the measure of association.
- Confounding (confounding bias):* A distortion of the measure of association between an exposure and outcome due to the presence of 1 or more confounders.
- Information bias:* A distortion of the measure of association between an exposure and outcome due to measurement errors in or misclassification of the exposure, covariate, or outcome variables.
- Propensity score:* The probability of being exposed or treated conditional on an individual's observed baseline characteristics. Methods using propensity scores (matching, stratification, weighting, covariate adjustment) can be used to reduce effects of confounding when analyzing observational data.
- Selection bias:* A distortion of the measure of association between an exposure and outcome that occurs due to the procedures used to select individuals for the study or the analysis.

cigarette users who resembled e-cigarette users in their observed baseline characteristics, leaving only 237 pairs of patients in the analysis.

The resulting analysis comparing e-cigarette users with nonusers is susceptible to various forms of selection bias (5). Selection bias due to loss to follow-up can occur if the risk for attrition differed by e-cigarette use and the probability of tobacco abstinence differed between persons who completed 3-month follow-up and those who did not. In this example, 317 of 1357 (23%) patients did not complete 3-month follow-up, primarily because they missed the assessment or withdrew from the study. In addition, 18 patients who completed the 3-month follow-up did not respond to the e-cigarette

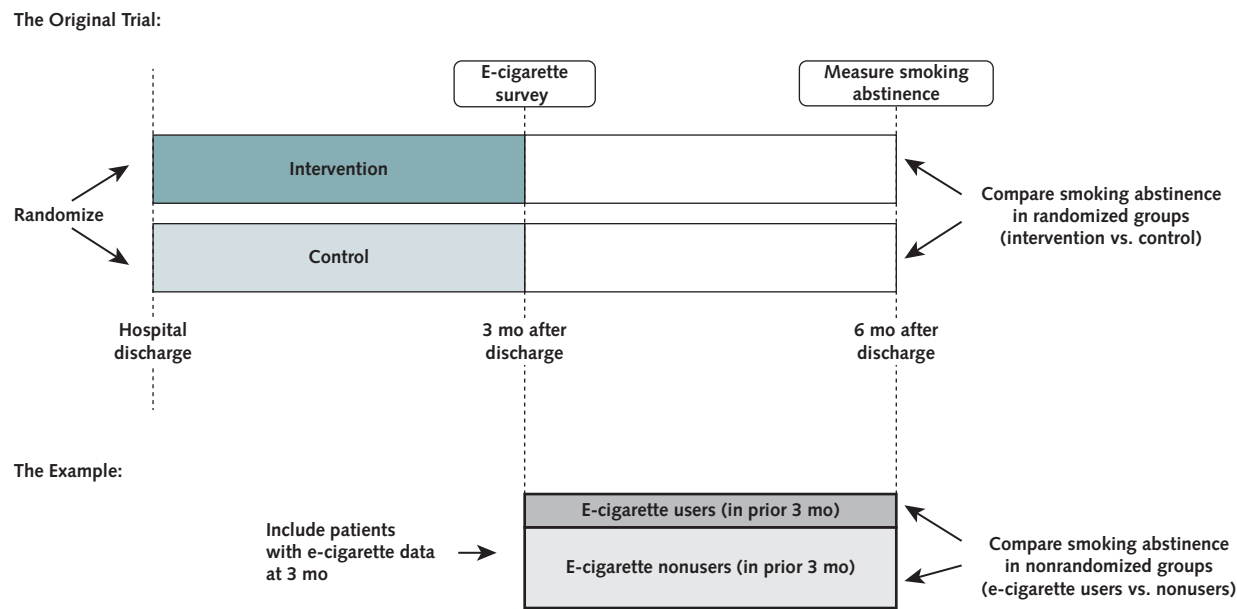
questions and were excluded from the analysis. Nonresponse can also create bias in an observational analysis. Finally, matching as done by the researchers means that the analysis sample will resemble only e-cigarette users in the sample rather than all patients. The resulting estimates answer the question of the effect of e-cigarette use among patients like the e-cigarette users, rather than among all smokers if they were to become e-cigarette users.

HOW DOES THE EXPOSURE COMPARE WITH THE RANDOMIZED INTERVENTION IN THE ORIGINAL TRIAL? HOW MIGHT THESE DIFFERENCES CREATE BIAS?

In the original trial, the intervention was sustained care versus standard care for smoking cessation. In contrast, the primary exposure used in the e-cigarette analysis was any self-reported use of e-cigarettes during the first 3 months after discharge.

Comparisons of exposures that differ from the randomized intervention in the original trial can be susceptible to confounding (5). In the example, the exposure (e-cigarette use) was not randomized. As a result, e-cigarette users may have differed from nonusers in more ways than just e-cigarette use. It is unlikely that all characteristics differing between users and nonusers were measured, allowing for the possibility of unmeasured confounding.

Figure. Design of original trial and analysis of nonrandomized exposure.



In the original Helping HAND 2 trial, patients were randomly assigned at hospital discharge to receive a 3-month smoking cessation intervention—either sustained or standard care. The primary outcome was smoking abstinence at 6 months after discharge. In the example, the researchers conducted a post hoc analysis of these data (the e-cigarette analysis) to determine whether e-cigarette use in the 3 months after hospital discharge was associated with subsequent tobacco abstinence. The researchers used all randomly assigned patients who remained in the study at 3 months after discharge and who completed an e-cigarette survey (75% of the initial trial population) to identify 237 propensity score-matched pairs with 1 exposed and 1 unexposed patient. Using these matched pairs, smoking abstinence at 6 months was compared between those who reported e-cigarette use in the 3 months after discharge (exposed) and those who did not (unexposed). Similar to analyses of observational data, this analysis needed to account for potential selection bias, confounding, and measurement error.

For example, suppose that e-cigarette use is related to age, with users being younger than nonusers. If age is also associated with successful tobacco cessation, it can confound the relationship between e-cigarette use and tobacco abstinence. In addition, the exposure in the example was self-reported and subject to measurement error. If e-cigarette use had been the randomized intervention, the device, nicotine dose, and frequency of use may have been monitored during follow-up.

Of note, the e-cigarette analysis and primary trial analysis shared the same outcome variable: tobacco abstinence at 6 months. Had outcomes for the 2 analyses differed, researchers would have had to consider how the outcome in the e-cigarette analysis was measured and the potential for information bias.

HOW DOES THE METHOD OF ANALYSIS CONTROL FOR THESE SOURCES OF BIAS?

In the example, researchers used propensity score matching to control for potential confounding. This analysis attempts to emulate some aspects of a randomized trial by pairing e-cigarette users and nonusers with similar baseline characteristics (6, 7). In particular, the researchers used demographic characteristics, smoking patterns, alcohol and marijuana use, and hospital discharge diagnosis to create a model that estimated the propensity (that is, the probability) of each person using e-cigarettes. This process produced a propensity score that was used along with the randomized study group from the original trial to match each e-cigarette user to a nonuser. This sample of propensity score-matched pairs was then used in the analyses.

If such matching is successful and the comparison groups are well balanced on observed baseline characteristics, analyses of matched pairs control for confounding due to observed factors (6, 7). Researchers should provide summaries of participant characteristics by exposure group before and after matching, with standardized differences, to demonstrate balance after matching. The number of adequate matches, or the degree of similarity in the distribution of propensity scores in the exposed and unexposed groups, also needs to be examined. When the exposed and unexposed groups differ greatly, comparisons can be unreliable (7). In addition, matching can result in a sample that is not representative of the population of interest (7). In the example, results apply to smokers like the e-cigarette users and may not be generalizable to all smokers.

Importantly, propensity score matching does not address potential confounding due to unmeasured factors. To address that possibility, the researchers conducted sensitivity analyses to demonstrate that their results could be explained away only by a strong, unmeasured confounder (8).

HOW ARE RESULTS FROM THE ANALYSIS INTERPRETED?

In discussing study results, researchers in the example emphasized that e-cigarette use was a self-selected and nonrandomized exposure, the analysis was subject to confounding, and the results may not

adequately reflect the underlying causal relationships. In addition, they noted lack of information about the type of e-cigarette device, frequency or duration of use, and timing of e-cigarette use relative to the randomized intervention as limitations. Finally, the researchers also called for further investigation of the association, including conduct of randomized controlled trials.

CONCLUSION

Post hoc analyses of randomized trial data can help answer important questions by making use of valuable data collected in controlled settings (9). Using trial data to compare nonrandomized exposures requires care because the comparisons are not protected by randomization. The questions outlined here can help readers understand the potential biases that need to be controlled for and considered when evaluating study results.

From American College of Physicians, Philadelphia, Pennsylvania (C.B.S., A.R.M.); University of Washington School of Medicine, Seattle, Washington (J.M.L.); and Johns Hopkins Bloomberg School of Public Health and Johns Hopkins School of Medicine, Baltimore, Maryland (E.G.).

Disclosures: Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M20-0071.

Corresponding Author: Anne R. Meibohm, PhD, American College of Physicians, 190 N. Independence Mall West, Philadelphia, PA 19106; e-mail, ameibohm@acponline.org.

Current author addresses and author contributions are available at Annals.org.

Ann Intern Med.

References

- Hernán MA, Hernández-Díaz S, Robins JM. Randomized trials analyzed as observational studies. *Ann Intern Med.* 2013;159:560-2. [PMID: 24018844] doi:10.7326/0003-4819-159-8-201310150-00709
- Rigotti NA, Tindle HA, Regan S, et al. A post-discharge smoking cessation intervention for hospital patients: Helping HAND 2 randomized clinical trial. *Am J Prev Med.* 2016;51:597-608. [PMID: 27647060] doi:10.1016/j.amepre.2016.04.005
- Rigotti NA, Chang Y, Tindle HA, et al. Association of e-cigarette use with smoking cessation among smokers who plan to quit after a hospitalization. A prospective study. *Ann Intern Med.* 2018;168:613-20. [PMID: 29582077] doi:10.7326/M17-2048
- Porta M. *A Dictionary of Epidemiology*. 6th ed. New York: Oxford Univ Pr; 2014.
- Ananth CV, Lavery JA. Biases in secondary analyses of randomized trials: recognition, correction, and implications. *BJOG.* 2016;123:1056-9. doi:10.1111/1471-0528.13732
- Rubin DB, Waterman RP. Estimating the causal effects of marketing interventions using propensity score methodology. *Stat Sci.* 2006;21:206-22. doi:10.1214/08833423060000000259
- Stuart EA. Matching methods for causal inference: a review and a look forward. *Stat Sci.* 2010;25:1-21. [PMID: 20871802]
- VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. *Ann Intern Med.* 2017;167:268-74. [PMID: 28693043] doi:10.7326/M16-2607
- Howard G, Howard VJ. Observational epidemiology within randomized clinical trials: getting a lot for (almost) nothing. *Prog Cardiovasc Dis.* 2012;54:367-71. [PMID: 22226006] doi:10.1016/j.pcad.2011.08.003

Current Author Addresses: Dr. Stack: Jazz Pharmaceuticals, 2005 Market Street, Suite 2100, Philadelphia, PA 19103.
Dr. Meibohm: American College of Physicians, 190 N. Independence Mall West, Philadelphia, PA 19106.
Dr. Liao: University of Washington School of Medicine, 1959 NE Pacific Street, Seattle, WA 98115.
Dr. Guallar: Welch Center for Prevention, Epidemiology and Clinical Research, Johns Hopkins Bloomberg School of Public Health, Room 2-645, Baltimore, MD 21205.

Author Contributions: Conception and design: C.B. Stack, A.R. Meibohm, J.M. Liao.
Drafting of the article: C.B. Stack, A.R. Meibohm, J.M. Liao.
Critical revision of the article for important intellectual content: C.B. Stack, A.R. Meibohm, J.M. Liao, E. Guallar.
Final approval of the article: C.B. Stack, A.R. Meibohm, J.M. Liao, E. Guallar.
Statistical expertise: C.B. Stack, A.R. Meibohm, E. Guallar.