Annals Understanding Clinical Research: Implications of Missing Data Due to Dropout

Joshua M. Liao, MD, and Catharine B. Stack, PhD

Through a series of "Understanding Clinical Research" articles, *Annals* aims to help nonstatisticians assess the implications of analytic methods used in research. These articles use illustrations from published studies to identify questions that are critical to understanding particular methods and appropriate interpretation of findings. This installment addresses the issue of missing data due to dropout (for the definition of this and other terms used in the article, see the Glossary)—a common problem that can affect the validity of trial findings. Several analytic methods exist to handle this problem; each has assumptions that require explicit statement and consideration (see the Table).

The Example

A randomized, controlled trial sought to determine whether adding spinal manipulative therapy to home exercise and advice would reduce short- and long-term back-related leg pain (1). The study randomly assigned 192 patients to either spinal manipulation with home exercise or home exercise alone over 12 weeks. Spinal manipulation therapy included up to 20 visits with a chiropractor, and home exercise included 4 visits to teach positioning, stabilization exercises, and posture awareness.

The primary outcomes were patient-reported back-related leg pain at 12 and 52 weeks, measured using an 11-point scale. The researchers found that spinal manipulation provided a clinically important advantage over home exercise at 12 weeks (between-group difference, −1.0 point [95% CI, −1.9 to −0.2 point]) but not at 52 weeks (between-group difference, −0.7 point [CI, −1.5 to 0.2 point]). Most of the trial's prespecified secondary outcomes also improved with spinal manipulation at 12 weeks, but only global improvement, satisfaction, and medication use showed sustained improvement at 52 weeks.

How Much Dropout Occurred? Why and When Did It Happen? In Which Study Group?

To understand the implications of dropout on results from randomized trials, readers should carefully examine the extent, timing, and sources of dropout. In this study, 1 participant in the home exercise group was lost to follow-up and did not provide outcomes at 12 weeks. Thirteen participants (8 in the home exercise group and 5 in the combined therapy group) were lost to follow-up and did not provide outcomes at 52 weeks. Most dropped out between their week-12 and week-26 visits. Reasons for dropouts were unclear.

How Were the Missing Outcomes Handled in the Analysis? What Assumptions Were Made?

The researchers analyzed data using a mixed-effects regression model, a method that utilizes all observed data. The model provides valid results as long as missing data follow a missing-at-random framework (2–6). Data are missing-at-random if they depend on other information collected during the trial, such as patient characteristics, previous outcomes, or other assessments. Despite its potentially confusing name, "missing-at-random" means that any systematic differences between missing and observed data can be explained by other information collected during the trial. The assumption that missingness is conditional on known values is then used in regression models to predict missing outcomes.

One alternative for handling missing data might have been to exclude the 13 participants with missing 52-week outcomes from the analysis. This kind of complete-case analysis assumes that data are missing-completely-at-random (2–6). This assumption holds when there is no relationship between the likelihood of missing and observed or unobserved data (that is, missing information results from a purely random process and missing and observed data have similar distributions). Common examples of this include technician error or machine malfunction.

Another tempting but flawed alternative would have been to "fill in" missing 52-week pain scores with a single value (such as a patient's pain score at baseline or at their last visit before dropping out, or the mean pain scores of patients who did not drop out). This application of single imputation makes strong assumptions that the 52-week pain scores are known with certainty in those who dropped out. It can lead to bias in either direction and always understates variance, with *P* values that are too small and CIs that are too narrow (2–6).

Were the Assumptions About the Missing Data Clinically Plausible?

Because the missing-at-random framework assumes that the tendency of data to be missing does not depend on missing values, analyzing trial results under this assumption is reasonable when little is known about the reasons for dropout. If the researchers had instead made the unlikely assumption that the data were missing-completely-at-random and performed a complete-case analysis, they could have biased their results. Similarly, single imputation would have

This article was published at Annals.org on 28 February 2017.

Glossary

Complete-case analysis: Analysis where participants with missing outcomes are excluded. Also referred to as a *completers analysis*.

Dropout: In a clinical trial, refers to a situation where a participant does not attend a scheduled visit and all subsequent study visits. Also referred to as *withdrawal*, *attrition*, or *loss to follow-up*.

Informative missing: Missing data that arise from a missing-not-at-random mechanism.

Missing-at-random (MAR): Mechanism for missing data whereby missingness is related to other observed data and the missing values can be predicted on the basis of what has been observed.

Missing-completely-at-random (MCAR): Mechanism for missing data whereby missingness is purely random and is not related to any observed or missing data.

Missing-not-at-random (MNAR): Mechanism for missing data whereby missingness is nonrandom and informative of the future outcomes.

Mixed-effects model: Type of regression model that properly handles repeated measurements from participants over time, includes all observed data, and does not require complete data from all participants.

Multiple imputation: Method for handling missing data where an imputation model is used to find imputed values for the missing items to produce a "complete" data set; the imputation is repeated multiple times, producing many imputed data sets, and each data set is analyzed; and appropriate procedures are then used to combine results in order to find estimates that have correct variance estimates. Multiple imputation may be done under missing-at-random or missing-not-at-random scenarios.

Noninformative missing: Missing data that arise from a missing-completely-at-random or missing-at-random mechanism.

Single imputation: Ad hoc methods for handling missing data where missing values are assumed to equal fixed values (such as the baseline value, the last value observed, or the mean for the treatment group) and the "complete" data set is then analyzed. Estimates found using single imputation methods are subject to bias and have variance estimates that are too low because the analysis treats the fixed values as if they had been observed.

overstated statistical significance and imposed very specific—and likely incorrect—assumptions about the missing outcomes.

However, mixed-effects regression models are not a panacea. In this study, for example, suppose that the patients who dropped out before 52 weeks had lower pain scores than those who did not drop out, and that their scores at 52 weeks could not be anticipated on the basis of prior measurements or observed covariates. In such a scenario, missing data would be informative because the likelihood of dropout is associated with the future outcome. The unknown reasons for dropout (for example, effective pain reduction from spinal manipulation) would provide critical information for estimating the true treatment response that is not accounted for by a missing-at-random assumption. When dropout is informative, missing data follow a missing-not-at-random mechanism; comparison of pain scores using mixed-effects regression models in that situation could be biased (2–6).

Unfortunately, there is no way to test whether data are missing-at-random or missing-not-at-random. Instead, researchers should explore the effect of different assumptions about missing data by conducting sensitivity analyses under various informative missingness scenarios (2–7).

Table. Key Questions for Assessing the Reporting and Handling of Missing Data Due to Dropout

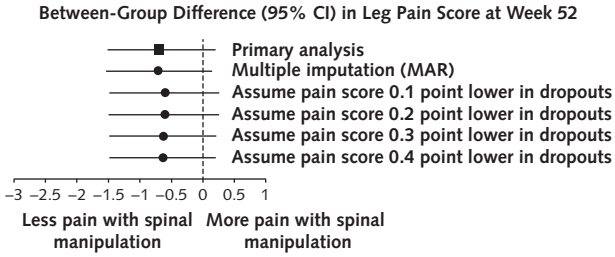
How much dropout occurred? Why and when did it happen? In which study group?
How were the missing outcomes handled in the analysis? What assumptions were made?
Were the assumptions about the missing data clinically plausible?
Were sensitivity analyses done to show how robust results were to changes in assumptions about the missing data? If so, did they target meaningful scenarios for the missing data?

Were Sensitivity Analyses Done to Show How Robust Results Were to Changes in Assumptions About the Missing Data? If So, Did They Target Meaningful Scenarios for the Missing Data?

The researchers reported 5 sensitivity analyses (Figure). They chose to perform the sensitivity analyses using multiple imputation (8), one of many possible approaches (2, 4–7). For each analysis, they first used observed data and applied a missing-at-random assumption to create 20 sets of imputed data, which they analyzed to find a set of 20 results. They then combined the results to find appropriate estimates of variance. This kind of multiple imputation approach avoids the problem of overstated statistical significance that exists with single imputation. The results from this multiple imputation model were very close to those from the mixed-effects regression model because both assumed data were missing at random.

Next, given the clinical plausibility that patients with less leg pain were more likely to drop out than those who remained in the study, the researchers performed 4 additional sensitivity analyses. These analyses used multiple imputation to generate an imputed score before subtracting a range of small but plausible amounts (0.1, 0.2, 0.3, and 0.4 point) from the 11-point pain score. Across the spectrum of these assumptions, results were similar to those from the primary analysis, suggesting that the results were robust to a set of

Figure. Plot of results from the primary and sensitivity analyses in Table 1 and Appendix Table 1 of Bronfort and colleagues' study (1).



The primary analysis used a mixed-effects model under an MAR assumption. Sensitivity analyses were conducted using multiple imputation under the following assumptions: missing outcomes were missing at random, and missing pain scores were 0.1, 0.2, 0.3, or 0.4 point lower than expected under an MAR framework. MAR = missing-at-random.

assumptions about informative missing outcomes (Figure).

CONCLUSION

In practice, every trial presents unique challenges with regard to missing data due to dropout, and no single method can be universally applied to every study. However, an approach that begins with an analysis based on a missing-at-random assumption about missing data, with additional sensitivity analyses to explore the effect of clinically plausible deviations from that assumption, is generally advised. The questions outlined here can help clinicians and peer reviewers assess how researchers have handled and reported missing data and their possible effect on trial results.

From University of Pennsylvania and American College of Physicians, Philadelphia, Pennsylvania.

Acknowledgment: The authors thank the *Annals* senior clinical and statistical editors for their input and review of earlier drafts of the manuscript.

Disclosures: Authors have disclosed no conflicts of interest. Forms can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M17-0195.

Requests for Single Reprints: Joshua M. Liao, MD, Department of Medicine, University of Pennsylvania, 3400 Spruce Street, Philadelphia, PA 19104.

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

Ann Intern Med. 2017;166:596-598. doi:10.7326/M17-0195

References

1. **Bronfort G, Hondras MA, Schulz CA, Evans RL, Long CR, Grimm R.** Spinal manipulation and home exercise with advice for subacute and chronic back-related leg pain: a trial with adaptive allocation. Ann Intern Med. 2014;161:381-91. [PMID: 25222385] doi:10.7326/M14-0006

2. **National Research Council, Division of Behavioral and Social Sciences and Education, Committee on National Statistics, Panel on Handling Missing Data in Clinical Trials.** The Prevention and Treatment of Missing Data in Clinical Trials. Washington, DC: National Academies Pr; 2010.

3. **Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al.** The prevention and treatment of missing data in clinical trials. N Engl J Med. 2012;367:1355-60. [PMID: 23034025] doi:10.1056/NEJMsr1203730

4. **Molenberghs G, Kenward M.** Missing Data in Clinical Studies. London: J Wiley; 2007.

5. **Carpenter JR, Kenward MG.** Missing data in randomised controlled trials - a practical guide. Publication RM03/JH17/MK. Birmingham, United Kingdom: National Institute for Health Research; 2008. Accessed at www.missingdata.org.uk on 25 January 2017.

6. **Bell ML, Fairclough DL.** Practical and statistical issues in missing data for longitudinal patient-reported outcomes. Stat Methods Med Res. 2014;23:440-59. [PMID: 23427225] doi:10.1177/0962280213476378

7. **Mallinckrodt C, Roger J, Chuang-Stein C, Molenberghs G, O'Kelly M, Ratitch B, et al.** Recent developments in the prevention and treatment of missing data. Ther Innov Regul Sci. 2014;48:68-80. doi:10.1177/2168479013501310

8. **White IR, Royston P, Wood AM.** Multiple imputation using chained equations: issues and guidance for practice. Stat Med. 2011;30:377-99. [PMID: 21225900] doi:10.1002/sim.4067

Current Author Addresses: Dr. Liao: Department of Medicine, University of Pennsylvania, 3400 Spruce Street, Philadelphia, PA 19104.
Dr. Stack: American College of Physicians, 190 N. Independence Mall West, Philadelphia, PA 19106.

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