

SPECIAL REPORT

The Prevention and Treatment of Missing Data in Clinical Trials

Roderick J. Little, Ph.D., Ralph D'Agostino, Ph.D., Michael L. Cohen, Ph.D., Kay Dickersin, Ph.D., Scott S. Emerson, M.D., Ph.D., John T. Farrar, M.D., Ph.D., Constantine Frangakis, Ph.D., Joseph W. Hogan, Sc.D., Geert Molenberghs, Ph.D., Susan A. Murphy, Ph.D., James D. Neaton, Ph.D., Andrea Rotnitzky, Ph.D., Daniel Scharfstein, Sc.D., Weichung J. Shih, Ph.D., Jay P. Siegel, M.D., and Hal Stern, Ph.D.

BACKGROUND

Missing data have seriously compromised inferences from clinical trials, yet the topic has received little attention in the clinical-trial community.¹ Existing regulatory guidances²⁻⁴ on the design, conduct, and analysis of clinical trials have little specific advice on how to address the problem of missing data. A recent National Research Council (NRC) report⁵ on the topic seeks to address this gap, and this article summarizes some of the main findings and recommendations of that report. The authors of this article served on the panel that prepared the report.

Missing data have seriously compromised inferences from clinical trials.¹ For example, editorials in the *Journal* have noted how missing data have limited the ability to draw definitive conclusions from weight-loss trials⁶ or could lead to incorrect inferences about drug safety.⁷ High rates of missing data that can affect conclusions occur in trials of treatments for many diseases.⁸⁻¹³ Since existing regulatory guidances²⁻⁴ lack specificity, in 2008 the Food and Drug Administration (FDA) requested that the NRC convene an expert panel to prepare “a report with recommendations that would be useful for FDA’s development of guidance for clinical trials on appropriate study designs and follow-up methods to reduce missing data and on appropriate statistical methods to address missing data for analysis of results.” This article summarizes some of the main findings and recommendations of the report⁵ of that panel. More details are provided elsewhere.¹⁴

The report focused primarily on phase 3 confirmatory clinical trials for assessing the safety and efficacy of drugs, biologic products, and

some medical devices, for which the bar of scientific rigor is set high. The use of randomized study-group assignments predominates in such studies, since this design feature ensures comparability of study groups and allows assessment of causation. However, many of the recommendations are applicable to early-phase randomized trials and epidemiologic studies in general.

Missing data are defined as values that are not available and that would be meaningful for analysis if they were observed. For example, measures of quality of life are usually not meaningful for patients who have died and hence would not be considered as missing data under this definition. We focus on missing outcome data here, though analysis methods have also been developed to handle missing covariates and auxiliary data.

KEY FINDINGS

Substantial instances of missing data are a serious problem that undermines the scientific credibility of causal conclusions from clinical trials. The assumption that analysis methods can compensate for such missing data are not justified, so aspects of trial design that limit the likelihood of missing data should be an important objective. In addition to specific aspects of trial design, many components of clinical-trial conduct can limit the extent of missing data. Finally, in studies with missing data, analysis methods that are based on plausible scientific assumptions should be used. For example, this consideration often rules out simple fixes, such as imputation by the last observation carried forward.¹⁰ Although there are better analysis alternatives to that method, they all require unverifiable assumptions. Thus,

sensitivity analyses should be conducted to assess the robustness of findings to plausible alternative assumptions about the missing data.

We now consider a number of specific missing-data issues that are intended to be representative and informative rather than comprehensive.

FOLLOW-UP AFTER TREATMENT DISCONTINUATION

A major source of missing data in clinical trials is participants who discontinue the assigned treatment because of adverse events, lack of tolerability, lack of efficacy, or simple inconvenience. Too many investigators incorrectly equate treatment discontinuation with study dropout; that is, outcomes are not recorded for participants who discontinue treatment. However, enrollees committed to participating in the study, not just to receiving the assigned treatment. When a study treatment is discontinued, efforts should be made to obtain the participant's consent for the collection of data on treatments and outcomes. When such efforts are successful, gathering these data after treatment discontinuation preserves the ability to analyze end points for all participants who underwent randomization and thus to make possible intention-to-treat inferences, which are grounded in randomization. It also allows exploration of whether the assigned therapy affected the use and efficacy of subsequent therapies and provides the ability to monitor side effects that might occur or persist after the discontinuation of treatment.⁷ The consensus of

the panel was that in many studies, the benefits of collecting outcomes after participants have discontinued treatment outweigh the costs.⁵

TRIAL DESIGN

Since there is no foolproof way to analyze data in the face of substantial amounts of missing data, we emphasize the role of design and trial conduct to limit the effect of missing data on regulatory decisions. Good clinical-trial design should clearly define the target population, along with efficacy and safety outcomes, and the likely effect of missing data should factor into decisions about reasonable alternative choices. The report⁵ states that “investigators, sponsors, and regulators should design clinical trials consistent with the goal of maximizing the number of participants who are maintained on the protocol-specified intervention until the outcome data are collected.”

Design elements for clinical trials can help to prevent missing data by reducing the number of participants for whom primary end-point data will be missing. A variety of design ideas are discussed in the report, and eight of them are shown in Table 1. Their relevance varies greatly according to setting, and they may have limitations or drawbacks that need to be considered.

An important and relatively neglected design issue is how to account for the loss of power from missing data in statistical inferences such as hypothesis tests or confidence intervals. The most common approach simply inflates the re-

Table 1. Eight Ideas for Limiting Missing Data in the Design of Clinical Trials.

Target a population that is not adequately served by current treatments and hence has an incentive to remain in the study.
Include a run-in period in which all patients are assigned to the active treatment, after which only those who tolerated and adhered to the therapy undergo randomization.
Allow a flexible treatment regimen that accommodates individual differences in efficacy and side effects in order to reduce the dropout rate because of a lack of efficacy or tolerability.
Consider add-on designs, in which a study treatment is added to an existing treatment, typically with a different mechanism of action known to be effective in previous studies.
Shorten the follow-up period for the primary outcome.
Allow the use of rescue medications that are designated as components of a treatment regimen in the study protocol.
For assessment of long-term efficacy (which is associated with an increased dropout rate), consider a randomized withdrawal design, in which only participants who have already received a study treatment without dropping out undergo randomization to continue to receive the treatment or switch to placebo.
Avoid outcome measures that are likely to lead to substantial missing data. In some cases, it may be appropriate to consider the time until the use of a rescue treatment as an outcome measure or the discontinuation of a study treatment as a form of treatment failure.

quired sample size in the absence of missing data to achieve the same sample size under the anticipated dropout rate, estimated from similar trials. This approach is generally flawed, since inflating the sample size accounts for a reduction in precision of the study from missing data but does not account for bias that results when the missing data differ in substantive ways from the observed data. In the extreme case in which the amount of bias from missing data is similar to or greater than the anticipated size of the treatment effect, detection of the true treatment effect is unlikely, regardless of the sample size, and the study is noninformative. When performing power calculations, one should consider sample-size computations for an intention-to-treat analysis that uses a hypothesized population treatment effect that is attenuated because of the inability of some study participants to adhere to the treatment. Alternatively, one could develop power analyses for statistical procedures that explicitly account for missing data and its associated uncertainty, as discussed below.

TRIAL PLANNING AND CONDUCT

The incidence of missing data varies greatly across clinical trials. Some of this variation is context-specific, but in many cases more careful attention to limiting missing data in trial planning and conduct can substantially reduce the problem. Eight practical ideas are listed in Table 2. These and other ideas are discussed in more detail in the panel report.^{5,14}

ANALYSIS METHODS

FOUR TYPES OF ADJUSTMENT APPROACHES

There is no universal method for handling missing data in a clinical trial, since each trial has its own set of design and measurement characteristics. The range of approaches to modeling and inference is extremely broad, and no single method or class of methods is suitable for all situations. The panel distinguished four different types of adjustment methods for missing data: complete-case analysis, single imputation methods, estimating-equation methods, and methods based on a statistical model.

In complete-case analysis, participants with missing data are simply excluded from the analysis.

In simple imputation methods, a single value is filled in for each missing value by means of methods such as the last observation carried forward and the baseline observation carried forward.

In estimating-equation methods, complete cases are weighted by the inverse of an estimate of the probability of being observed. For example, the probability of an outcome being observed might be modeled with the use of baseline data, and then the complete cases might be weighted by the inverse of their estimated probabilities of being observed. More generally, model-based estimates are augmented with weighted residuals from the model.

As an example of a statistical model, continuous repeated measures might be assumed to have a normal distribution with a specified form

Table 2. Eight Ideas for Limiting Missing Data in the Conduct of Clinical Trials.

Select investigators who have a good track record with respect to enrolling and following participants and collecting complete data in previous trials.
Set acceptable target rates for missing data and monitor the progress of the trial with respect to these targets.
Provide monetary and nonmonetary incentives to investigators and participants for completeness of data collection, as long as they meet rigorous ethical requirements. ^{15,16}
Limit the burden and inconvenience of data collection on the participants, and make the study experience as positive as possible.
Provide continued access to effective treatments after the trial, before treatment approval.
Train investigators and study staff that keeping participants in the trial until the end is important, regardless of whether they continue to receive the assigned treatment. Convey this information to study participants.
Collect information from participants regarding the likelihood that they will drop out, and use this information to attempt to reduce the incidence of dropout.
Keep contact information for participants up to date.

of mean and covariance matrix. Methods that are based on a statistical model such as this include maximum likelihood, in which estimates and standard errors are based on the likelihood function given the observed data; Bayesian methods, in which inferences are based on a statistical model that includes an assumed prior distribution for the measurements; and multiple imputation, in which multiple sets of plausible values for missing data are created from their model-based predictive distribution, and estimates and standard errors are obtained with the use of multiple-imputation combining rules.¹⁷ Hogan et al.¹⁸ provide a relatively nontechnical description of these methods, with examples.

THREE SCENARIOS FOR MISSING DATA

The properties of these methods depend on the mechanisms leading to missing data. A useful taxonomy^{5,19} distinguishes methods that assume one of three scenarios. In the first scenario, data are missing completely at random, which implies that the missing data are unrelated to the study variables. In particular, the complete cases are representative of all the original cases as randomized. In the second scenario, data are missing at random, which implies that recorded characteristics can account for differences in the distribution of missing variables for observed and missing cases. In the third scenario, data are missing not at random, which implies that recorded characteristics do not account for differences in the distribution of the missing variables for observed and missing cases.

For example, in an imaginary clinical trial, some participants have had improvement or worsening of their condition, and some have had toxic effects and have dropped out before the final outcome could be recorded. The assumption that data are missing completely at random presumes that outcomes for those who dropped out would be expected to be similar to outcomes for participants who did not drop out, so the data from dropouts can be ignored without bias. The less stringent assumption that data are missing at random implies that outcomes for participants who dropped out would be expected to be similar to outcomes for participants who did not drop out with similar baseline characteristics and similar intermediate measures up to that time, so missing outcomes can be modeled on the basis of outcomes of similar

participants who did not drop out. Only the assumption that the data are missing not at random allows for the possibility that events that were not observed (e.g., severe toxicity or disease progression occurring since the last visit) may have influenced the decision to drop out, and thus outcomes are likely to be different from those of similar participants who did not drop out. Models that are based on the assumption that the data are not missing at random must make further assumptions about the effect of such possibilities.

These assumptions about the missing-data mechanisms have an effect on the appropriateness of the different methods. We do not recommend using the complete-case-analysis approach to missing data, since it requires the unrealistic assumption that the data are missing completely at random. This method does not exploit partial information in incomplete cases, which could improve the estimates when handled properly.

Simple imputation methods such as the last observation carried forward and baseline observation carried forward are commonly applied, in part because they are simple and easy to understand, but they are overused. We do not recommend them, since their validity hinges on assumptions that are often unrealistic. For example, the last-observation-carried-forward method makes the assumption that the outcomes of participants do not change after they have dropped out, leading to biased treatment effects when this assumption is not justified.¹⁰ These methods, like many other methods that impute a single value for the missing data, do not propagate imputation uncertainty and thus yield inappropriately low estimates of standard errors and P values. Thus, we recommend that single-imputation methods, such as last observation carried forward and baseline observation carried forward, should not be used as the primary approach to the treatment of missing data “unless the assumptions that underlie such methods are scientifically justified.”⁵

FAVORED METHODS AND SENSITIVITY ANALYSES

In general, the panel favored estimating-equation methods and methods that are based on a statistical model for the data. In particular, weighted estimating equations and multiple-imputation models have an advantage in that they can be used to incorporate auxiliary information about

the missing data into the final analysis, and they give standard errors and P values that incorporate missing-data uncertainty. Analyses that are performed with such methods often assume that missing data are missing at random, and such an assumption often makes sense for the primary analysis. However, the observed data can never verify whether this assumption is correct. Therefore, to assess robustness, sensitivity analyses are recommended. We advocate sensitivity analyses that are easy to interpret by clinicians. In the report, we included examples of such analyses that are based on pattern-mixture and selection models. One pattern-mixture approach for comparing two treatments assumes that participants who drop out have a mean outcome that deviates from that of participants who do not drop out by an offset — which we term $d(1)$ for treatment 1 and $d(2)$ for treatment 2 — and then explores the effect on the findings of various choices of the offsets in the two study groups. If the treatment effect is qualitatively maintained for the range of offsets that are considered to be clinically plausible, then the findings are considered to be robust. Extensions apply the offsets to means adjusted for available covariates and modify the offsets to be suitable for categorical outcomes.

SIX PRINCIPLES FOR DRAWING INFERENCES

The following six principles for drawing inferences from incomplete data are applicable in a wide variety of settings.

First, determine if possible whether the values that are missing are meaningful for analysis and hence meet the definition of missing data.

Second, formulate an appropriate and well-defined causal primary measure of treatment effect in terms of the data that were intended to be collected. It is important to distinguish what is being estimated from the method of estimation, which may vary according to assumptions.

Third, as much as possible, document the reasons why data are missing. For example, did the patient move to get a better job or did he or she drop out of the trial because of a severe side effect? Knowing the reasons for missing data can help formulate sensible assumptions about observations that are missing. A related idea is to seek and collect auxiliary variables that may be predictive of the outcome and of dropping out, since analysis methods such as multiple im-

putation and weighted estimating equations can exploit these data to reduce bias from missing data and improve the precision of estimates.

Fourth, decide on a primary set of assumptions about the missing-data mechanism. In some cases, the primary assumptions can be that data are missing at random. Assumptions about the missing-data mechanism (as opposed to details of the computational algorithms) must be transparent and accessible to clinicians.

Fifth, conduct a statistically valid analysis under the primary missing-data assumptions.

And sixth, assess the robustness of inferences about treatment effects to various missing-data assumptions by conducting a sensitivity analysis that relates inferences to one or more parameters that capture departures from the primary missing-data assumption, such as the pattern-mixture analysis outlined above. Sensitivity analysis is a relatively new area, and further research on the best methods is needed. The interpretation of the collective results from a sensitivity analysis when some of the analyses are in opposition to the primary analysis also needs more consideration, though the panel's report contained some approaches to addressing such situations.

In summary, there is no easy fix for missing data at the analysis stage. Too many current analyses of clinical trials apply naive methods for missing-data adjustment that make unjustified assumptions, such as the last-observation-carried-forward approach. In reports and interpretations of trial results, the handling of missing data requires a scientifically defensible analysis coupled with a sensitivity analysis to assess robustness. No matter what approach is taken, there is no way to adequately test the robustness of the assumptions about missing data required by the analysis. This need to rely on untestable assumptions regarding missing data reinforces the importance of preventing missing data in the first place. The key is to design and carry out the trial in a way that limits the problem of missing data.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

From the Departments of Biostatistics (R.J.L.) and Statistics (S.A.M.), University of Michigan, Ann Arbor; the Department of Mathematics and Statistics, Boston University, Boston (R.D.); National Research Council of the National Academies, Washington, DC (M.L.C.); the Departments of Epidemiology (K.D.) and Biostatistics (C.F., D.S.), Johns Hopkins University, Baltimore; the Department of Biostatistics, University of Washing-

ton, Seattle (S.S.E.); the Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia (J.T.F.); the Center for Statistical Sciences, Program in Public Health, Brown University, Providence, RI (J.W.H.); the International Institute for Biostatistics and Statistical Bioinformatics, Universiteit Hasselt and Katholieke Universiteit, Leuven, Belgium (G.M.); the School of Public Health, University of Minnesota, Minneapolis (J.D.N.); Departamento de Economia, Universidad Torcuato Di Tella, Buenos Aires (A.R.); the Department of Biostatistics, University of Medicine and Dentistry of New Jersey School of Public Health, New Brunswick (W.J.S.); Johnson & Johnson, Radnor, PA (J.P.S.); and the Department of Statistics, University of California, Irvine (H.S.). Second and subsequent authors are listed alphabetically.

1. O'Neill RT, Temple R. The prevention and treatment of missing data in clinical trials: an FDA perspective on the importance of dealing with it. *Clin Pharmacol Ther* 2012;91:550-4.
2. Statistical principles for clinical trials; step 5: note for guidance on statistical principles for clinical trials: International Conference on Harmonisation, Topic E9. London: European Medicines Agency, 1998 (<http://www.ich.org/products/guidelines/efficacy/article/efficacy-guidelines.html>).
3. Committee for Medical Products for Human Use. Guideline on missing data in confirmatory clinical trials. London: European Medicines Evaluation Agency, April 2009 (<http://www.ema.europa.eu/pdfs/human/ewp/177699endraft.pdf>).
4. ICH harmonised tripartite guideline: choice of control group and related issues in clinical trials: International Conference on Harmonisation, Topic E10. London: European Medicines Agency, 2001 (<http://www.ema.europa.eu/pdfs/human/ich/036496en.pdf>).
5. National Research Council. The prevention and treatment of missing data in clinical trials. Washington, DC: National Academies Press, 2010 (http://www.nap.edu/catalog.php?record_id=12955).
6. Ware JH. Interpreting incomplete data in studies of diet and weight loss. *N Engl J Med* 2003;348:2136-7.
7. Lagakos SW. Time-to-event analyses for long-term treatments — the APPROVe Trial. *N Engl J Med* 2006;355:113-7.
8. Kemmler G, Hummer M, Widschwendter C, Fleischhacker

WW. Dropout rates in placebo-controlled and active-control clinical trials of antipsychotic drugs: a meta-analysis. *Arch Gen Psychiatry* 2005;62:1305-12.

9. Rabinowitz J, Levine SZ, Barkai O, Davidov O. Dropout rates in randomized clinical trials of antipsychotics: a meta-analysis comparing first- and second-generation drugs and an examination of the role of trial design features. *Schizophr Bull* 2009;35:775-88.
10. Molnar FJ, Man-Son-Hing M, Hutton B, Fergusson DA. Have last-observation-carried-forward analyses caused us to favour more toxic dementia therapies over less toxic alternatives? A systematic review. *Open Med* 2009;3(2):e31-e50.
11. Lurie I, Levine SZ. Meta-analysis of dropout rates in SSRIs versus placebo in randomized clinical trials of PTSD. *J Nerv Ment Dis* 2010;198:116-24.
12. Lipinski MJ, Cauten CA, Biondi-Zoccai GG, et al. Meta-analysis of randomized controlled trials of statins versus placebo in patients with heart failure. *Am J Cardiol* 2009;104:1708-16.
13. Raboud JM, Montaner JSG, Thorne A, Singer J, Schechter MT. Impact of missing data due to dropouts on estimates of the treatment effect in a randomized trial of antiretroviral therapy for HIV-infected individuals. *J Acquir Immune Defic Syndr Hum Retrovirol* 1996;12:46-55.
14. Little RJ, Cohen ML, Dickersin K, et al. The design and conduct of clinical trials to limit missing data. *Stat Med* 2012 July 25 (Epub ahead of print).
15. Emanuel EJ. Undue inducement: nonsense on stilts? *Am J Bioeth* 2005;5:9-13. [Erratum, *Am J Bioeth* 2006;6:54.]
16. Committee on Assessing the System for Protecting Human Research Participants. Responsible research: a systems approach to protecting research participants. Washington, DC: National Academy Press, 2002.
17. Little RJ, Rubin DB. Statistical inference with missing data. 2nd ed. New York: Wiley, 2002.
18. Hogan JW, Roy J, Korkontzelou C. Handling drop-out in longitudinal studies. *Stat Med* 2004;23:1455-97.
19. Rubin DB. Inference and missing data. *Biometrika* 1976;63:581-92.

DOI: 10.1056/NEJMSr1203730

Copyright © 2012 Massachusetts Medical Society.

CLINICAL PROBLEM-SOLVING SERIES

The *Journal* welcomes submissions of manuscripts for the Clinical Problem-Solving series. This regular feature considers the step-by-step process of clinical decision making. For more information, please see authors.nejm.org.