

## Addressing Missing Data in Clinical Trials

Thomas R. Fleming, PhD

The reliability and interpretability of results from clinical trials can be substantially reduced by missing data. Frequently used approaches to address these concerns, such as upward adjustments in sample sizes or simplistic methods for handling missing data, including last-observation-carried-forward, complete-case, or worst-case analyses, are usually inadequate. Although rational imputation methods may be useful to treat missingness after it has occurred, these methods depend on untestable assumptions. Thus, the preferred and often only satisfactory approach to addressing missing data is

to prevent it. Procedures should be in place to maximize the likelihood that outcome data will be obtained at scheduled times of evaluation for all surviving patients who have not withdrawn consent. To meaningfully reduce missing data, it is important to recognize and address many factors that commonly lead to higher levels of missingness.

*Ann Intern Med.* 2011;154:113-117.

For author affiliation, see end of text.

[www.annals.org](http://www.annals.org)

**T**he reliability and interpretability of results from randomized clinical trials is greatly influenced by the quality of trial conduct. A factor having a substantial effect on this quality is the amount of missing data, particularly when patients who are lost to follow-up have an inherently different level of frailty and prognosis. Such missingness induces risk for substantial bias.

A recent study illustrates the important differences between patients who have and those who do not have missing data and the resulting inadequacy of some simplistic approaches to handling missingness that are popular, particularly in regulatory settings (1). The study described changes from baseline in 6-minute walk distance over 48 months in patients with pulmonary arterial hypertension who were receiving sildenafil. In the **Figure**, at each scheduled visit, it is clear that patients who are about to discontinue follow-up are less vigorous. A complete case estimate of 6-minute walk distance, defined by averaging the outcome of patients who remain under observation, would provide a significant overestimate of the true long-term performance of the original cohort of patients in the trial because patients remaining under follow-up tend to have a more favorable clinical course. A last-observation-carried-forward analysis also would provide an overestimate because patients who discontinue are likely to have continued decline in stamina, and some will die soon after discontinuation of follow-up. Imputing the worst possible outcome for those with missing data would be justified for patients who have died (2) and may be appropriate for those too ill to complete the 6-minute walk distance assessment but would underestimate the clinical course for others who do not have a rapid decline when they are no longer under observation.

A preferred approach to addressing missing data is to prevent it. In order to preserve the integrity of randomization, all patients should be followed until the complete capture of trial outcomes (3–5), even after patients have discontinued randomized treatment or initiated other interventions. Achieving such follow-up enables the conduct of a proper “as-randomized” analysis in which the study outcome is assessed in all patients (6). This analysis evaluates an intervention as part of an experimental regimen that also includes effects of ancillary care and rescue ther-

apies that might be provided to patients and addresses the questions of greatest relevance to real-world settings because of its unconditional nature (7). In contrast, a “per-protocol” analysis that includes only patients who have adhered closely to what is prescribed in the study protocol does not maintain the comparability of treatment groups that randomization provides because it excludes randomly assigned participants on the basis of outcomes after baseline. It also addresses only a conditional situation that would not validly apply to the entire set of eligible patients, diminishing its relevance and interpretability in the real-world setting (8). Per-protocol analyses of safety outcomes that truncate follow-up soon after treatment discontinuation also are problematic, especially about the evaluation of treatment-induced risks that could arise on a delayed basis.

Procedures should be in place to maximize the likelihood that outcome data will be obtained at scheduled times of evaluation for all surviving patients who have not withdrawn consent. This article discusses approaches to the design, conduct, and analysis of clinical trials that will enhance the reliability of results by reducing missingness.

### ADDRESSING FACTORS IN TRIAL DESIGN THAT INFLUENCE LEVELS OF MISSING DATA

The level of missing data can be reduced by creative approaches in the formulation of protocols, informed consents, and data collection forms, as well as in the selection and education of both investigators and patients.

First, there should be proper distinction in study protocols between reasons for nonadherence (that is, for not receiving randomized therapy and hence for being “off study treatment”) versus nonretention (that is, for not ob-

See also:

#### Print

Key Summary Points ..... 114

#### Web-Only

Conversion of graphics into slides

**Key Summary Points**

Often the only satisfactory approach to addressing missing data is to prevent it.

It is important to pursue many approaches to reduce the occurrence of missing data.

Protocols should more clearly distinguish between reasons for taking a patient "off study treatment" (nonadherence) vs. "off study" (nonretention).

Follow-up should not be discontinued because of inappropriate characterization of "withdrawal of consent."

The informed consent process should more clearly alert patients to the negative effect that incomplete capture of outcomes has on trial integrity and credibility.

Protocol-specified increases in sample sizes to address missing data should be recognized to simply produce more precise biased estimates.

Protocols should specify performance standards for achieving high quality of trial conduct, including high levels of data capture.

Studies should involve only those investigators who are committed to follow all patients until death or capture of all trial outcomes, even if the patients have discontinued randomized treatment or initiated other interventions.

Creative and effective procedures should be implemented during enrollment and follow-up to enhance achieving prespecified targeted levels of retention.

An oversight process should be in place during trial conduct to ensure the achievement of performance standards, including targeted levels of data capture.

taining outcome information and hence for being "off study") (9). There are only 2 valid reasons a patient can be off study: withdrawal of consent or the achievement of all required efficacy and safety end point information. If study end points are properly defined, the occurrence of death satisfies the latter condition rather than inducing missing data. Unfortunately, it is common for protocols to provide a long list of reasons that the patient will be off study, such as inability to tolerate the intervention, toxicity, physician choice, or need for other therapies. These may be valid reasons for nonadherence (for being off study treatment), but not for nonretention (for being off study). The protocol should separately list the 2 reasons a patient could go off study and the many reasons the patient could go off study treatment, with an indication that all reasonable efforts should be made to ensure patients who go off study treatment be consistently followed for outcomes unless they have withdrawn consent. The terminology *inactive*

rather than *lost to follow-up* or *off study* has been suggested for missing patients who have not withdrawn consent to reflect the possibility of achieving resumption of adherence or recovery of missing data (10).

Second, the term *withdrawal of consent* should not be used simply because the patient no longer wishes to receive randomized treatment or actively continue to return for follow-up assessments or simply to justify why efforts are not being made to continue to follow some patients who have discontinued their randomized intervention (9). Rather, the term should be used only when the patient no longer wishes to participate in the trial and no longer authorizes the investigators to make efforts to continue to obtain their outcome data. Ideally, if patients withdraw their consent, it should be done in writing (11). The rates of properly validated withdrawal of consent usually should be in the range of 1%, although higher rates might be expected in antipsychotic settings or where patients have impaired cognitive function. It is important that investigators be educated and evaluated about the proper use of the term *withdrawal of consent*, and data monitoring committees should regularly assess whether the term was being used properly.

Third, patients should be adequately educated during the informed consent process about the continued scientific relevance of their data even if they discontinue treatment, as well as the deleterious effect that missing data has on trial integrity and credibility (8, 12–15). Patients join trials not only for their own personal gain but also for altruistic reasons. Although they should be informed that they can withdraw consent at any time, they also deserve to be informed that missing data diminish the scientific value of all patients' altruistic contributions. Wording to this effect has been successfully added to informed consent forms to enable patients to make more informed decisions about their willingness to join a trial and to participate in continued follow-up without feeling inappropriately pressured to do so (10).

Fourth, protocols should not give a misleading sense that biostatisticians have adequately corrected for the negative effect of substantial levels of missing data. Specifically, sections on statistical considerations frequently indicate that 10% to 50% increases in sample size have been made to address expected levels of missing data. Clarification is lacking that such adjustments in sample size address only the variability and not the bias induced by missing data so that these adjustments simply result in obtaining more precise biased estimates.

Fifth, protocols should specify the performance standards that should be met to achieve high quality of trial conduct, including specification of targeted levels of data capture (16). A performance standards document should be developed before finalization of the study protocol that specifies targets about the enrollment and eligibility rates, event rate, adherence and retention rates, and currentness of data capture.

Finally, forms and procedures for data collection should include creative approaches to reduce missing data (12). The burden on patients should be minimized by reducing the number of visits and assessments conducted, reducing the number of variables collected, formulating user-friendly case report forms, and enlarging the visit window. Investigators should be screened for their track record for retention and should be well trained in procedures to maximize data capture.

### IMPLEMENTING PROCEDURES DURING ENROLLMENT AND FOLLOW-UP THAT INFLUENCE MISSINGNESS

Creative and effective procedures should be implemented during enrollment and follow-up to enhance the likelihood of high levels of retention. Such procedures are illustrated by HIVNET (HIV Network for Prevention Trials) 012, which was designed to evaluate a regimen to reduce the risk for mother-to-child transmission of HIV in developing country settings (17). The HIV status of infants was to be assessed through 18 months after delivery. Many believed targeted levels of at least 95% retention could not be achieved in such settings because mother-infant pairs who were coming into Kampala, Uganda, to participate in the trial were very mobile and living “up country” without home addresses.

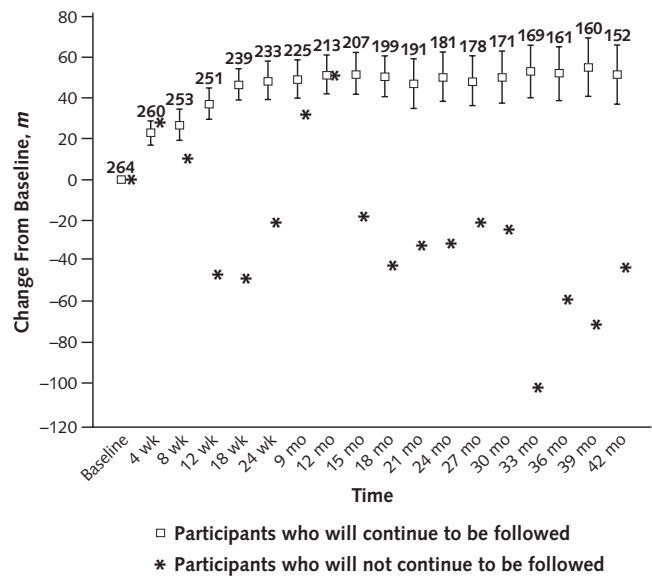
The HIVNET 012 research team recognized that the level of bias from missing data would not be lower simply because it would be more difficult to achieve retention. To reduce this risk for bias, the team formed a group of health visitors trained as health social workers, community health educators, or home visitors for maternal and child health, who had several important roles, including increasing retention (18).

The health visitors, assigned to participants by geographic area, increased retention by creating rapport, at first with the participants and then with their families. They provided education about primary health care components, such as nutrition, family planning, and immunizations. They also assured participants of their confidentiality and observed customer care principles, making participants comfortable, thanking them when they came for scheduled visits to the clinic, and demonstrating caring attitudes toward the sick.

The health visitors obtained locator information from the participants and additional information about contacts to assist in achieving effective follow-up. They made regular home visits to keep close contact, strengthen relationships, and provide reminders about scheduled visits. The health visitors recorded daily activities in report forms, held monthly meetings to evaluate their activities and solve problems, and attended regular meetings with the trial’s principal investigators and study coordinators to share updates on each participant.

The landmark HIVNET 012 trial established an economically and practically feasible intervention in a devel-

**Figure. Change from baseline in 6-minute walk distance over 48 months after initiation of treatment in patients with pulmonary arterial hypertension.**



At each scheduled visit, the average change from baseline in 6-minute walk distance is plotted separately for the subgroups that will and will not remain under follow-up at the time of the next scheduled visit. The numbers above the squares represent, at each scheduled visit, the number of patients in the subgroup that will remain under follow-up at the time of the next scheduled visit. Hence, these are the numbers of patients who contribute to the calculations of the means and the 95% CIs for that subgroup.

oping country setting to be safe and effective. Its integrity was substantially enhanced by high levels of retention. Contrary to claims that even short-term rates of loss to follow-up would be at least 20%, infant HIV status was captured in 97.1% of the 619 mother-infant pairs in the HIVNET 012 trial at the week 14 to 16 visit and in 95.5% at the month 18 visit (19).

### IMPLEMENTING OVERSIGHT PROCESSES DURING TRIAL CONDUCT TO ENHANCE RETENTION

Regular oversight by peer reviewers during a trial can improve the implementation of creative procedures for enhancing the quality of trial conduct, such as those procedures specified in the HIVNET 012 trial to improve retention. For illustration, the networks sponsored by the National Institute of Allergy and Infectious Diseases conducting HIV prevention trials have created multidisciplinary “study monitoring committees” that meet during the trial to conduct periodic reviews of data pooled across intervention groups. Their goal is to assess whether pre-specified targets for performance are met for measures, such as the enrollment and eligibility rates, event rate, adherence and retention rates, and currentness of data cap-

ture, and then to make recommendations about steps that could be implemented to improve these rates. The study monitoring committee has semi-independent membership that includes network researchers who do not have leadership responsibilities in the trials. Its findings are reported to the independent data monitoring committee, and both committees can assess the appropriateness of trial continuation if these findings reveal serious and uncorrected deficiencies in trial conduct. Substantial improvements in retention achieved in many trials having oversight by study monitoring committees have been among the important benefits of this oversight process.

### ILLUSTRATING CREATIVE APPROACHES TO REDUCE MISSING DATA AND ITS EFFECT ON TRIAL INTEGRITY

In psychotherapy research, clinical trials often have had high levels of missing data. This resulted from frequent nonadherence to study regimens and widely implemented practices of either excluding patients who did not complete the protocol-specified randomized treatments or discontinuing follow-up after randomized treatments were stopped. Creative approaches were implemented in the evaluation of treatments for posttraumatic stress disorder to effectively reduce missingness and thus enhance the reliability and relevance of results to real-world clinical practice (8, 20). Trials were designed by using a clinically relevant control regimen that represented a true version of standard of care; reducing restrictions about patient access to ancillary care during the trial; defining relatively nonrestrictive eligibility criteria; allowing enrollment of patients from diverse clinical settings; evaluating a range of outcome measures beyond the target symptoms assessed in the primary end point; and educating patients and investigators about the importance of continuing follow-up, even after treatment discontinuation or initiation of rescue interventions. In one posttraumatic stress disorder trial, the level of missingness that would have been greater than 30% was reduced to less than 10% by using the procedures mentioned (8). In the same study with high levels of retention, the primary analysis indicated lack of benefit on any outcome measure, in contrast to a misleading indication of benefit when analysis was restricted to patients who completed treatment.

### APPROACHES TO AVOID

Some approaches for reducing the risk for missing data should be avoided. Changing the definition of a primary end point to reduce the risk for missing data would be inappropriate if such a change meaningfully compromises the end point's clinical relevance. For example, reducing the follow-up to obtain a reduction in patients with missing data should not be encouraged in chronic disease settings when it is important to understand the longer-term benefit–risk profile of the experimental intervention. Fur-

thermore, the clinical relevance of a time-to-event end point based on the composite of events, “progression of major symptoms,” and “death” often would be meaningfully compromised by forming a broader composite that also includes the event “discontinuation of treatment” or “exposure to rescue treatment.” Compromising clinical relevance is an unacceptable price to pay to eliminate the risk for informative missingness that occurs through failure to follow patients after they have discontinued randomized interventions.

### SUMMARY

Missing data on outcome measures in clinical trials meaningfully reduce the integrity and interpretability of results, adversely affecting science and thus the medical community pursuing an evidence-based practice of medicine. The effect is substantial when a nontrivial fraction of patients have missing data on primary and secondary efficacy and safety end points in the trial and when mechanisms for missingness are related to or informative about the outcome measures. For too long, there has been overreliance on methods to adjust for missing data in clinical trials. Worse yet, these methods often have been naively inadequate, such as making adjustments in sample size that address the variability but not the bias induced by missingness or using simplistic approaches, such as last-observation-carried-forward, complete-case, or worst-case analyses.

Although rational imputation methods may be useful to treat missingness after it has occurred, all such methods depend on untestable assumptions. The preferred and often only satisfactory approach to addressing missing data is to prevent it. Among the procedures that should be in place to reduce missing data are training of investigators about the importance of retention (9, 12) and obtaining signed investigator agreements indicating commitments to continue follow-up efforts even after patients have discontinued randomized treatment or initiated other interventions. Consideration should be given to discontinuing sites at which investigators are unable or unwilling to achieve this level of data capture. The Key Summary Points and the Institute of Medicine's extensive report about handling missing data (12) provide additional approaches to reducing missingness by addressing the many factors that commonly contribute to missing data. There is wisdom in the recognition that “an ounce of prevention is worth a pound of cure.”

From the University of Washington, Seattle, Washington.

**Note:** This article was delivered on 9 September 2009 as an invited presentation at the National Academy of Sciences Workshop on the Handling of Missing Data in Clinical Trials.

**Acknowledgment:** The author thanks Laura Guay for insights about retention procedures in the HIVNET 012 trial, Scott Emerson for his

important insights shared during many discussions about this topic, and Gary Layton and Mike Oakes for providing access to their creative analysis of the pulmonary arterial hypertension data.

**Grant Support:** By the National Institutes of Health/National Institute of Allergy and Infectious Disease (R37 AI 29168).

**Potential Conflicts of Interest:** Disclosures can be viewed at [www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M10-1856](http://www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M10-1856).

**Requests for Single Reprints:** Thomas R. Fleming, PhD, Department of Biostatistics, University of Washington, Box 357232, Seattle, WA 98195-7232; e-mail, [tfleming@u.washington.edu](mailto:tfleming@u.washington.edu).

Author contributions are available at [www.annals.org](http://www.annals.org).

## References

- Galiè N, Ghofrani HA, Torbicki A, Barst RJ, Rubin LJ, Badesch D, et al; Sildenafil Use in Pulmonary Arterial Hypertension (SUPER) Study Group. Sildenafil citrate therapy for pulmonary arterial hypertension. *N Engl J Med*. 2005;353:2148-57. [PMID: 16291984]
- Lubsen J, Kirwan BA. Combined endpoints: can we use them? *Stat Med*. 2002;21:2959-70. [PMID: 12325112]
- Lavori PW. Clinical trials in psychiatry: should protocol deviation censor patient data? *Neuropsychopharmacology*. 1992;6:39-48. [PMID: 1571068]
- Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. *Br J Cancer*. 1977;35:1-39. [PMID: 831755]
- Pocock S. *Clinical Trials: A Practical Approach*. New York: J Wiley; 1983.
- Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c332. [PMID: 20332509]
- Fleming TR, Rothmann MD, Lu HL. Issues in using progression-free survival when evaluating oncology products. *J Clin Oncol*. 2009;27:2874-80. [PMID: 19414672]
- Schnurr PP, Friedman MJ, Foy DW, Shea MT, Hsieh FY, Lavori PW, et al. Randomized trial of trauma-focused group therapy for posttraumatic stress disorder: results from a department of veterans affairs cooperative study. *Arch Gen Psychiatry*. 2003;60:481-9. [PMID: 12742869]
- Wittes J. Missing inaction: preventing missing outcome data in randomized clinical trials. *J Biopharm Stat*. 2009;19:957-68. [PMID: 20183458]
- Lachin JM. Statistical considerations in the intent-to-treat principle. *Control Clin Trials*. 2000;21:167-89. [PMID: 10822117]
- Cleland JG, Torp-Pedersen C, Coletta AP, Lammiman MJ. A method to reduce loss to follow-up in clinical trials: informed, withdrawal of consent. *Eur J Heart Fail*. 2004;6:1-2. [PMID: 15012911]
- National Research Council. *The Prevention and Treatment of Missing Data in Clinical Trials*. Panel on Handling Missing Data in Clinical Trials. Committee on National Statistics, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Pr; 2010.
- Levine RJ. *Ethics and Regulation in Clinical Research*. 2nd ed. Baltimore: Urban & Schwarzenberg; 1986.
- Gordon EJ, Prohaska TR. The ethics of withdrawal from study participation. *Accountability in Research*. 2006;13:285-309.
- Faden RR, Beauchamp TL. *A History and Theory of Informed Consent*. New York: Oxford Univ Pr; 1986.
- Fleming TR. Identifying and addressing safety signals in clinical trials [Editorial]. *N Engl J Med*. 2008;359:1400-2. [PMID: 18768938]
- Guay LA, Musoke P, Fleming T, Bagenda D, Allen M, Nakabiito C, et al. Intrapartum and neonatal single-dose nevirapine compared with zidovudine for prevention of mother-to-child transmission of HIV-1 in Kampala, Uganda: HIVNET 012 randomised trial. *Lancet*. 1999;354:795-802. [PMID: 10485720]
- Katongole SNR, Kalyesubula I, Kikonyogo F, Lubega D, Musoke MP, Guay L. Use of home health visiting system to minimize loss to follow up in Ugandan clinical trials. Presented at the Fourteenth International AIDS Conference, Barcelona, Spain, 7-12 July 2002. Abstract MoPeE3726.
- Jackson JB, Musoke P, Fleming T, Guay LA, Bagenda D, Allen M, et al. Intrapartum and neonatal single-dose nevirapine compared with zidovudine for prevention of mother-to-child transmission of HIV-1 in Kampala, Uganda: 18-month follow-up of the HIVNET 012 randomised trial. *Lancet*. 2003;362:859-68. [PMID: 13678973]
- Schnurr PP, Friedman MJ, Engel CC, Foa EB, Shea MT, Chow BK, et al. Cognitive behavioral therapy for posttraumatic stress disorder in women: a randomized controlled trial. *JAMA*. 2007;297:820-30. [PMID: 17327524]

**Author Contributions:** Drafting of the article: T.R. Fleming.  
Final approval of the article: T.R. Fleming.  
Statistical expertise: T.R. Fleming.  
Obtaining of funding: T.R. Fleming.