Randomized trials with missing outcome data: how to analyze and what to report

Rolf H.H. Groenwold MD PhD, Karel G.M. Moons PhD, Jan P. Vandenbroucke MD PhD

The gold-standard study design to evaluate the effects of medical treatment is the randomized trial. Assignment to treatment groups is a random process, and baseline differences in prognostic factors are due to chance. Consequently, at baseline, both groups are expected to be statistically comparable with respect to both known and unknown prognostic factors for the outcomes being studied. To maintain this expected comparability, it is common wisdom that participant data should be analyzed according to the assigned treatment group (i.e., intention-to-treat principle). However, there may be loss to follow-up (i.e., patients for whom the outcome is not known).

This problem was discussed in depth by a panel from the National Research Council[1] and was recently commented upon.[2,3] The report lists a number of precautions that can be taken to limit losses to follow-up. Nevertheless, missing outcomes are almost inevitable (e.g., if patients do not return for follow-up appointments) regardless of precautions, and loss of outcome information can amount to 50%.[4] In this article, we aim to explain and illustrate the main problem of missing outcomes in the analyses of randomized trials, potential solutions, and what should be reported.

The problem

Irrespective of whether a study is a randomized trial or an observational follow-up, a major problem in both designs is patients who are lost to follow-up.[5,6] Patients who drop out may do so for reasons that are linked to their prognosis. If the characteristics of the patients who drop out differ between treatment groups, patients who remain in the study may no longer be comparable for their prognosis, and an incorrect estimate of treatment effect may result.[4,7,8] This holds true even if a perfect balance of prognostic factors was achieved at baseline and intention-to-treat analysis is performed.[4] An intention-to-treat analysis merely indicates that participant data are analyzed according to the treatment to which they were assigned. The term intention-to-treat holds no information about how missing out-

comes were handled in the analysis, and participants with missing outcomes are typically omitted from the analysis. This results in a "complete case intention-to-treat analysis."

An example of bias due to missing outcomes

We illustrate with a simple example how missing outcome data can result in biased estimates of treatment effects (Table 1). Suppose that a placebo-controlled randomized trial involving 2000 participants was conducted to evaluate the effect of an active treatment on a certain outcome. Because of randomization, there was an equal number of men and women in each treatment group. Sex is a risk factor for the outcome: the risk is three times higher in men than in women (30% v. 10%, respectively, in the placebo group). If outcome data are available for all patients (scenario 1), a protective treatment effect is observed (risk ratio [RR] 0.80).

Scenario 2 in Table 1 shows an extreme example of missing outcome data, in which all women in the active treatment group are lost to follow-up (e.g., because they are more prone to adverse drug reactions). As a result, the risk of an event in the treatment group is overestimated (24% instead of 16%), because only individuals at high risk of the outcome remain. Consequently, the observed treatment effect is biased (RR 1.20), indicating a harmful effect of treatment, even though the treatment is actually protective. If the opposite situation occurred (i.e., all

---

KEY POINTS

- In the event of missing outcome data in a randomized trial, analyzing data for only participants for whom outcomes were observed may bias the estimate of treatment effect.

- A comparison of prognostic characteristics between study groups limited to participants included in the analysis may help to identify selective missingness of outcome data.

- Contingency plans to deal with missing outcome data by controlling for baseline characteristics, multiple imputation or inverse probability weighting should be prespecified in the trial protocol.

---

men in the treatment group were lost to follow-up), the observed treatment effect would be severely biased (RR 0.40).

This example shows that the bias due to missing outcome data may lead to an overestimation or underestimation of the treatment effect. Instead of removing from the analysis all women in the active treatment group for whom no outcome was observed, we could easily implement the (unrealistic) assumption that either all or none of the women in this group experienced the outcome. This would result in biased estimates of the treatment effect as well (RR 3.10 and RR 0.60, respectively). However, it is unlikely that anyone would consider this assumption to be reasonable.

Scenario 3 presents a more realistic situation that may nevertheless be more puzzling. In this scenario, we assume that 25% of people in each group have missing outcome data. At first glance, such a finding would make most readers think that there would be no problem, because missingness is the same in both groups. However, in the treatment group, the losses are among the women, and in the placebo group the losses are among the men. Because the baseline risk of the outcome is different for men and women, the risk of the outcome (as estimated using the available data) in the treatment and placebo groups will be biased. The observed treatment effect will also be biased (RR 1.12).

Even though sex was perfectly balanced between groups at baseline, for scenarios 1 and 2, including only participants in the analysis for whom the outcome was observed resulted in an imbalance in sex between the groups. In other words, even though an intention-to-treat analysis was performed, the estimated treatment effect was biased because of an imbalance in a prognostic variable (sex, in this example), which was the result of missing outcome data.

When does missing outcome data lead to biased effect estimates in an intention-to-treat analysis?

If outcome data are missing, they are typically expected to be selectively missing because prognosis often determines if patients will react to treatment (or placebo) or experience adverse effects. In scenarios 2 and 3 in Table 1, missingness of the outcome data was related to sex and to treatment group. If missingness of the outcome data is somehow related, directly or indirectly, to prognostic characteristics at baseline as well as treatment group, this will create a baseline imbalance in prognosis among those with observed outcomes. Note that this bias is not necessarily conservative and is not remedied by intention-to-treat analysis.

Missing outcome data due to nonprognostic factors are inconsequential (i.e., this will not bias the estimate of the treatment effect).[9] If missingness of the outcome depends on treatment status but is independent of prognostic factors (e.g., a random 50% of participants in the treatment

	Treatment group, no. (%) of events		Estimated treatment effect,* RR	What to report: characteristics of participants included in the analysis	
Scenario	Treatment	Placebo		Treatment, no. (%) of participants	Placebo, no. (%) of participants
Scenario 1: no missing outcomes					
Women	40/500 (8)	50/500 (10)		500/1000 (50)	500/1000 (50)
Men	120/500 (24)	150/500 (30)		500/1000 (50)	500/1000 (50)
Total	160/1000 (16)	200/1000 (20)	0.80		
Scenario 2: 25% missing outcomes (all women from treatment group)					
Women	–	50/500 (10)		0/500 (0)	500/1000 (50)
Men	120/500 (24)	150/500 (30)		500/500 (100)	500/1000 (50)
Total	120/500 (24)	200/1000 (20)	1.20		
Scenario 3: 25% missing outcomes that affect both treatment groups in a different way					
Women	20/250 (8)	50/500 (10)		250/750 (33)	500/750 (67)
Men	120/500 (24)	75/250 (30)		500/750 (67)	250/750 (33)
Total	140/750 (19)	125/750 (17)	1.12		

Table 1: Examples of bias due to missing outcomes in a randomized trial and how to report on trials with missing outcomes

Note: RR = risk ratio.
*In all scenarios, the estimated treatment effect is unbiased if stratified by sex. In scenario 2, the estimated treatment effect (RR) for men is 0.8 ([120/500]/[150/500]); the RR for women cannot be estimated. In scenario 3, the RR for men and women is the same (RR 0.8; men: RR = [120/500]/[75/250]; women: RR = [20/250]/[50/500]).

group are lost to follow-up), this will also not bias the estimate of the treatment effect.

What to report

The report by the National Research Council clearly states that researchers should describe the reason why patients were lost to follow-up and, hence, why outcomes were missing (e.g., participants moved away, which may signify a good prognosis).[1] We would like to add that, to make the process transparent to readers, a second table should be included, showing the distribution of baseline characteristics among the treatment groups for patients for whom outcomes were observed and who were included in the analysis.

The usual guidelines on the reporting of randomized trials clearly indicate that a "baseline table" of prognostic characteristics should be presented, to permit the reader to judge whether these characteristics were indeed balanced between the treatment groups immediately after randomization.[10] For example, this would mean including a table showing that the percentage of women among the treatment groups was indeed balanced at baseline (50% v. 50%). To identify imbalances in prognostic characteristics due to missing outcomes, authors should also present a "second baseline table"; that is, a comparison of

prognostic baseline characteristics between the participants in the study groups who were actually included in the analysis (i.e., those with observed outcomes) .

In scenario 2 presented in Table 1, this second table would indicate that the percentage of women is no longer balanced between groups (treatment: 0%; placebo: 50%). In scenario 3, this table would show that the percentage of women is imbalanced between groups (treatment: 33%; placebo: 67%).

A frequently used method to identify the potential for bias due to missing data is to compare participants with and without missing values. In scenario 2, such a table would indeed show that the percentages of women and participants in the treatment group are imbalanced between those with and without missing outcome data (of those with missing outcomes, 100% are women and 100% are in the treatment group; of those without missing outcomes, 33% are women and 33% are in the treatment group), indicating a potential for bias because of missing outcome data. In scenario 3, however, this table would show that the percentage of women (missing: 50%; not missing: 50%) and participants in the treatment group (missing: 50%; not missing: 50%) is equal among those with and without missing outcomes. Hence, a table comparing participants with and without missing values

Table 2: Guidance on how to analyze and report missing outcome data* in randomized trials		
Extent of missing outcome data	**How to analyze**	**How to report**
Small	• Main analysis should be intention-to-treat, including participants for whom the outcome was observed • Consider performing sensitivity analysis to quantify the effect of missing outcome data (e.g. using multiple imputation, regression adjustment or inverse probability weighting) on study results	• Baseline characteristics of the participants who were randomized to each study group • Proportion of missing outcomes per group • Characteristics of participants for whom no outcomes were observed • (Possible) reasons for missing outcome data • Emphasis on results of the main analysis; results of sensitivity analyses can be discussed in an appendix
Extensive	• Main analysis might rely on analytical methods that handle missing outcome data (e.g., multiple imputation, regression adjustment or inverse probability weighting), extending an intention-to-treat analysis	• Baseline characteristics of participants who were assigned to each study group • Proportion of missing outcomes per group • Characteristics of participants for whom no outcomes were observed • (Possible) reasons for missing outcome data • Characteristics of participants who were assigned to each study group and who were included in the analysis; this might replace the usual table of baseline characteristics • Emphasis on the results of the main intention-to-treat analysis, in which analytical methods to handle missing outcome data were used
*One of the drivers of bias due to missing outcome data is the proportion of missing outcome data in relation to the number of events. However, any cut-off is arbitrary. Even less than 5% (notably in the case of rare events) missing outcome data may result in considerable bias if missingness of the outcome is related to prognostic characteristics as well as to treatment. One way to assess the effect of missing outcome data is to use analytical methods that handle missing data and discuss any differences.		

would incorrectly reassure researchers and readers that the missing outcomes did not potentially result in bias, even though it did. If space is limited, a supplement on the journal's website might be a suitable place to report the proposed second baseline table.

Other issues concerning missing data that are important to report include what efforts were made in the design of the study to prevent missing data, the extent of missing data, how missing data were handled in the analysis, and an explicit statement about the potential for bias due to missing data. These issue have been put forward by others.[2,3,5,11–13] Guidance on the reporting and analysis of trials with missing outcomes is summarized in Table 2. We refer readers to the literature for further information on these topics.[5,9,11,14,15]

How to analyze

One way of dealing with missing data in medical research is to impute (i.e., fill in) the missing values.[5] There is essentially no difference between imputing a missing baseline covariate or a missing outcome value.[5] Nevertheless, researchers may feel uncomfortable when it comes to imputation of the outcome.

The most straightforward way to deal with imbalances due to selective missingness of the outcome in a randomized trial is to control for the imbalanced prognostic characteristics just as one would do in an observational study.[8,16,17] One might hold the view that, in the presence of missing outcomes, a randomized trial becomes an observational therapeutic study, in which treatment groups typically differ in baseline prognosis because of confounding. In observational epidemiologic studies, differences in prognostic characteristics between treatment groups are also usually adjusted for in the analysis (confounding adjustment).[18] The advantage of this approach is its simplicity and familiarity. Moreover, it becomes immediately clear to readers what happens: the authors performed an adjustment for prognostic imbalances, which were not due to randomization but occurred after randomization and probably not by chance. In all scenarios presented in Table 1, adjustment (e.g., stratification) for sex would indeed yield an unbiased treatment effect estimates (RR 0.8).

In the report by the National Research Council,[1] more intricate statistical methods (multiple imputation and inverse probability weighting) are discussed to control for the bias due to missing outcome data.[5,7,19] These methods have theoretical advantages above the simple adjustment method, because they do not rely on a model for adjustment but instead "recreate" in some sense the original study population. They may even use observed post-randomization variables or secondary outcomes to impute missing data for the primary outcome.[8] Hence, multiple imputation and inverse probability weighting are more flexible than conventional adjustment methods to address bias due to missing outcomes in randomized trials.[5,7,8,19] Multiple imputation can even be used in cases with both missing baseline and outcome data.

If the adjustment model, the multiple imputation model and the inverse probability weighting model are correctly specified and the same variables are included in each model, all will yield the same results in terms of bias and precision.[8] Even if multiple imputation or inverse probability weighting is preferred, it remains useful, as a first step, to assess the impact of missing outcomes by the more simple adjustment analysis using complete cases; in this situation, the same variables used for multiple imputation or inverse probability weighting should be used for adjustment.

Regardless of the method chosen, however, one should be aware that even after adjustment, multiple imputations or inverse probability weighting, the situation of the missing outcomes is only remedied to the extent that their missingness can be accounted for by the observed prognostic variables. This is clearly spelled out in the report by the National Research Council,[1] and leads the authors of the report to state that all such analytic solutions should only be seen as forms of sensitivity analysis.[2] Missing outcomes may still result in a bias, if missingness depends on unobserved prognostic variables or on the unobserved value of the outcome itself.[19] However, given that such bias only came into existence after randomization and reasonable comparability may have existed at baseline, bias may be less severe than in an observational study of the same treatment, and such bias may be largely remedied by adjustment or imputation based on the known prognostic baseline characteristics.

Conclusion

The key strength of randomized trials is that random allocation of participants implies that the treatment groups are expected to be comparable with respect to all (observed and unobserved) prognostic characteristics at baseline. To maintain this expectation of baseline comparability of groups, randomized trials are routinely analyzed according to the intention-to-treat principle. Despite all precautions, missing outcome data will inevitably be an issue in all randomized trials, and analyzing data for only participants for whom outcomes were observed may imbalance prognostic variables if the missingness of the out-

come data is related to prognostic characteristics as well as to treatment. In such cases, the baseline balance that randomization strives for will be distorted, and the analysis and inferences will actually approach that of an observational study. A trial with very large, even infinite, sample size will not help. Nor will an intention-to-treat analysis; instead, it may actually give false reassurance if there was indeed no imbalance at baseline.

Because the most important known prognostic characteristics are usually observed in a trial, researchers can easily check whether missing outcomes have led to bias by comparing prognostic baseline characteristics between the study groups, and limiting this comparison to patients with observed outcomes and who were thus included in the analysis. Any imbalance in prognostic characteristics identified using this "second baseline table" can be controlled for in the analysis, as one would do in an observational study, or by use of more flexible methods such as multiple imputation or inverse probability weighting.

Given that missing outcome data are inevitable in any randomized trial, researchers should be prepared to deal with this situation. Contingency plans to deal with missing outcome data should be prespecified in the trial protocol, because standard intention-to-treat analyses may be biased in the event of missing outcomes.

References

1. National Research Council. *The prevention and treatment of missing data in clinical trials*. Washington: National Academies Press, 2010. Available: www.nap.edu/catalog/12955.html (accessed 2012 Nov. 3).
2. Little RJ, D'Agostino R, Cohen ML, et al. The prevention and treatment of missing data in clinical trials. *N Engl J Med* 2012;367: 1355-60.
3. Ware JH, Harrington D, Hunter DJ, et al. Missing Data. *N Engl J Med* 2012;367:1353-4.
4. Wood AM, White IR, Thompson SG. Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. *Clin Trials* 2004;1: 368-76.
5. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393.
6. Altman DG. Missing outcomes in randomized trials: addressing the dilemma. *Open Med.* 2009;3:e51-3.
7. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;15:615-25.
8. Groenwold RHH, Donders ART, Roes KCB, et al. Dealing with missing outcome data in randomized trials and observational studies. *Am J Epidemiol* 2012;175:210-7.
9. Westreich D. Berkson's bias, selection bias, and missing data. *Epidemiology* 2012;23:159-64.
10. Schulz KF, Altman DG, Moher D. CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Ann Intern Med* 2010;152:726-32.
11. van Buuren S. *Flexible Imputation of Missing Data*. Boca Raton (FL): Chapman & Hall/CRC; 2012.
12. Higgins JPT, Altman DG, Sterne JAC; Cochrane Statistical Methods Group and the Cochrane Bias Methods Group. Assessing risk of bias in included studies. In: Higgins JPT, Green S, editors. *Cochrane handbook for systematic reviews of interventions. Version 5.1.0 [updated March 2011].* Oxford (UK): Cochrane Collaboration; 2011. Available: handbook.cochrane .org/ (accessed 2014 Feb. 1).
13. Higgins JP, Altman DG, Gøtzsche PC, et al.; Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928.
14. Donders AR, van der Heijden GJ, Stijnen T, et al. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol* 2006;59:1087-91.
15. Daniel RM, Kenward MG, Cousens SN, et al. Using causal diagrams to guide analysis in missing data problems. *Stat Methods Med Res* 2012;21:243-56.
16. Mallinckrodt CH, Lane PW, Schnell D, et al. Recommendations for the primary analysis of continuous endpoints in longitudinal clinical trials. *Drug Inf J* 2008;42:303-19.
17. Baker SG, Fitzmaurice GM, Freedman LS, et al. Simple adjustments for randomized trials with nonrandomly missing or censored outcomes arising from informative covariates. *Biostatistics* 2006;7:29-40.
18. Klungel OH, Martens EP, Psaty BM, et al. Methods to assess intended effects of drug treatment in observational studies are reviewed. *J Clin Epidemiol* 2004;57:1223-31.
19. Little RJA, Rubin DB. *Statistical analysis with missing data. 2nd edition*. New York (NY): Wiley-Interscience; 2002.

Affiliations: Julius Center for Health Sciences and Primary Care (Groenwold, Moons), University Medical Center, Utrecht, the Netherlands; Department of Clinical Epidemiology (Vandenbroucke), Leiden University Medical Center, Leiden, the Netherlands.