Addressing missing outcome data in meta-analysis

Dimitris Mavridis,1,2 Anna Chaimani,1 Orestis Efthimiou,1 Stefan Leucht,3,4 Georgia Salanti1

1Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece
2Department of Primary Education, University of Ioannina, Ioannina, Greece
3Department of Psychiatry and Psychotherapy, Technische Universität München, München, Germany
4Department of Psychiatry, University of Oxford, Oxford, UK

Correspondence to Dimitris Mavridis, dimi.mavridis@gmail.com

INTRODUCTION

The term attrition is widely used in the clinical trials literature to refer to situations where outcome data are not available for some participants. Missing data may invalidate results from clinical trials by reducing precision and potentially biasing the results. Systematic reviewers often assume that the missing outcome problem has been resolved at the trial level. However, in many clinical trials a complete case analysis or suboptimal imputation techniques are employed and the problem is accumulated in a quantitative synthesis of trials via meta-analysis. The risk of bias due to missing data depends on the missingness mechanism. Most statistical analyses assume missing data to be missing at random, which is an unverifiable assumption. The aim of this paper is to present methods used to account for missing outcome data in a systematic review and meta-analysis.

Methods

The following methods handle missing outcome data are presented: (1) complete cases analysis, (2) imputation methods from observed data, (3) best/worst case scenarios, (4) uncertainty interval for the summary estimate and (5) a statistical model that makes assumption about how treatment effects in missing data are connected to those in observed data. Examples are used to illustrate all the methods presented.

Results

Different methods yield different results. A complete case analysis leads to imprecise and potentially biased results. The best-case/worst-case scenarios give unrealistic estimates, while the uncertainty interval produces very conservative results. Imputation methods that replace missing data with values from the observed data do not properly account for the uncertainty introduced by the unobserved data and tend to underestimate SEs. Employing a statistical model that links treatment effects in missing and observed data, unlike the other methods, reduces the weight assigned to studies with large missing rates.

Conclusions

Unlike clinical trials, in systematic reviews and meta-analyses we cannot adapt pre-emptive methods to account for missing outcome data. There are statistical techniques implemented in commercial software (eg, STATA) that quantify the departure from the missing at random assumption and adjust results appropriately. A sensitivity analysis with increasingly stringent assumptions on how parameters in the unobserved and observed data are related is a sensible way to evaluate robustness of results.

ABSTRACT

Objective Missing outcome data are a common problem in clinical trials and systematic reviews, as it compromises inferences by reducing precision and potentially biasing the results. Systematic reviewers often assume that the missing outcome problem has been resolved at the trial level. However, in many clinical trials a complete case analysis or suboptimal imputation techniques are employed and the problem is accumulated in a quantitative synthesis of trials via meta-analysis. The risk of bias due to missing data depends on the missingness mechanism. Most statistical analyses assume missing data to be missing at random, which is an unverifiable assumption. The aim of this paper is to present methods used to account for missing outcome data in a systematic review and meta-analysis.

Methods

The following methods handle missing outcome data are presented: (1) complete cases analysis, (2) imputation methods from observed data, (3) best/worst case scenarios, (4) uncertainty interval for the summary estimate and (5) a statistical model that makes assumption about how treatment effects in missing data are connected to those in observed data. Examples are used to illustrate all the methods presented.

Results

Different methods yield different results. A complete case analysis leads to imprecise and potentially biased results. The best-case/worst-case scenarios give unrealistic estimates, while the uncertainty interval produces very conservative results. Imputation methods that replace missing data with values from the observed data do not properly account for the uncertainty introduced by the unobserved data and tend to underestimate SEs. Employing a statistical model that links treatment effects in missing and observed data, unlike the other methods, reduces the weight assigned to studies with large missing rates.

Conclusions

Unlike clinical trials, in systematic reviews and meta-analyses we cannot adapt pre-emptive methods to account for missing outcome data. There are statistical techniques implemented in commercial software (eg, STATA) that quantify the departure from the missing at random assumption and adjust results appropriately. A sensitivity analysis with increasingly stringent assumptions on how parameters in the unobserved and observed data are related is a sensible way to evaluate robustness of results.
III. Missing not at random (MNAR) or informatively missing (IM)

Even accounting for all the available observed information, the probability that an observation is missing still depends on the unseen observations themselves. Participants may dropout for reasons that are associated with the actual effect of the intervention. In schizophrenia trials, for example, placebo arms show larger dropout rate than patients treated with antipsychotics because of placebo’s lack of efficacy. Analysis of the participants who completed the study under MNAR would provide a biased estimate of the relative treatment effect. When missing data are MCAR or MAR they are termed ignorable. A MNAR mechanism is termed non-ignorable.

We use a hypothetical example to illustrate differences between the three categories. Consider an RCT with 200 participants randomised equally (1:1) to the experimental or control group (Table 1). We assume that the true response rate is 33.3% in the control group and 50% in the experimental group so that the estimated OR should be 2. In the MAR scenario, 10% of the participants dropped out because they missed the appointment. In the MAR scenario, young people dropped out because summer started and they left for vacation. In both these scenarios the missing rate across treatment groups is the same because groups are expected to have participants with similar baseline characteristics (eg, age). Therefore, the probability of dropping out is the same in both the groups. In the MAR scenario, 40% of the participants who did not see any improvement dropped the study. The number of those not improved is larger in the control group; hence, missing rate is larger in the control group.

Reasons for dropping out are related to the intervention received and more specifically to the actual outcome of the study. Missingness of dropping out refers to 27 and 20 participants, respectively.)

Estimons of missing outcome data in individual studies and their meta-analysis

Trials often employ methods to impute values for missing outcomes. Two very common approaches include (1) replacing missing values with the mean value of the participants who provided data (simple imputation) and (2) replacing missing values with the last observed value. The latter method is called last observation carried forward (LOCF) and is routinely used in mental health trials. The LOCF approach has been criticised for producing biased results, as conditions in mental health care are rarely stable and usually involve progressive conditions. Hence, it is not sensible to assume that participants who dropped out at intermediate steps would have remained stable until the end of the study (especially when data are MCAR the LOCF cannot be defended because participants could have left the study for reasons associated with their unobserved outcome data eg, they got worse). Simple imputation of missing values using the two aforementioned methods usually underestimates the SE for the outcome because it fails to account for the fact that missing values are imputed rather than observed. Multiple imputations replace missing values by a number of different plausible imputed values and subsequently combine them to get parameter estimates. Plausible imputed values are sampled from their predictive distributions given the observed data (eg, using regression models). The SE of the treatment effect is estimated by the within variance of each imputed dataset as well as from the variance between the imputed datasets. A naive or unprincipled imputation method may create more problems than it solves by introducing bias both in parameter estimates and their SE. When data are MAR, multiple imputation may correct for bias but it will still yield biased results when MNAR holds. To give unbiased results, the variable(s) that are predictive of missing data should be included in the imputation model. Published study reports typically present results together for fully observed and imputed outcomes. Consequently, a meta-analysis is not given much choice but to synthesise study outcomes as reported in trials, even when the imputation technique is inappropriate. In some cases, trial reports present the outcomes for completers only as well as the results from the merged sample of observed and imputed outcomes.

Synthesis of studies with missing outcome data

It is often the case that studies report results from the participants who provided the outcome of interest and they only report the number of participants for which the outcome is unknown. From a meta-analysis perspective several synthesis options exist and are outlined below.

### Table 1 Hypothetical example of missing data mechanisms applied to a randomised trial with 100 participants in each arm. We assume that the true response rate is 33.3% in the control group and 50% in the experimental group so that the estimated OR should be 2

<table>
<thead>
<tr>
<th>Missing data mechanism</th>
<th>MCAR</th>
<th>MAR</th>
<th>MNAR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>control group</td>
<td>experimental group</td>
<td>control group</td>
</tr>
<tr>
<td>Completers</td>
<td>90</td>
<td>90</td>
<td>60</td>
</tr>
<tr>
<td>Non-completers</td>
<td>10</td>
<td>10</td>
<td>40</td>
</tr>
<tr>
<td>Successes in completers</td>
<td>30</td>
<td>45</td>
<td>20</td>
</tr>
<tr>
<td>Reason for dropping out</td>
<td>10% of participants missed their appointment</td>
<td>Young people (40% of the total sample) dropped out because summer started and wanted to leave for vacation</td>
<td>40% of participants who did not see any improvement dropped out (roughly 67% in the control group and 50% in the experimental group did not see any improvement, hence on average the 40% who dropped out refers to 27 and 20 participants, respectively)</td>
</tr>
<tr>
<td>Estimated OR (using completers)</td>
<td>2.00</td>
<td>2.00</td>
<td>1.84</td>
</tr>
<tr>
<td>95% CI (1.10 to 3.65)</td>
<td>95% CI (0.96 to 4.18)</td>
<td>95% CI (0.97 to 3.50)</td>
<td></td>
</tr>
</tbody>
</table>
Complete cases meta-analysis
This is usually the reference approach in many meta-analyses. From each study, only individuals whose outcome is known are included. If the MCAR assumption holds then a complete cases meta-analysis will give unbiased results and the only consequence of missing data is the loss in power. The larger the missing rate, the less reliable the results of this analysis when the data are MNAR.10

Best-case and worst-case scenarios
A typical simple imputation technique for dichotomous outcomes includes the best-case and worst-case scenarios.14 The best-case scenario assumes that all missing participants have a favourable outcome in the experimental group and poor outcome in the control group; the converse is assumed for the worst-case scenario. These two extremes are typically used as a sensitivity analysis and may produce unrealistic results in practice, especially if missing rates are high.

Uncertainty intervals
Gamble and Hollis14 suggested that studies for which there is a big discrepancy between best-case and worst-case scenarios should be down-weighted. Best-worst case scenarios give an interval, called uncertainty interval, for the treatment effect that includes all uncertainty due to missing data.14 A pseudo SE is estimated from this for each study and is subsequently used to downweight studies. A summary estimate is computed using the revised weights from the uncertainty intervals. This method has low power when there is a large amount of missing data.14

Using a statistical model that relates missing data to observed data
For the case of missing binary outcomes, White and colleagues presented a meta-analysis model where the degree of departure from the MAR assumption is quantified by the informative missingness OR (IMOR).15 IMOR describes the relationship between the unknown odds of the outcome among missing participants and the known odds among observed participants.15 16 By relating missing treatment effects to the observed ones it is possible to adjust treatment effects within each study. The adjusted treatment effects are then summarised via meta-analysis. Higgins et al16 suggested a sensitivity analysis for the IMOR approach with risks imputed for the missing data over a plausible range of values. Expert opinion can be used to elicit information on how the risk in the missing participants is related to that of the observed participants and use that information to adjust treatment effects. Experts may be asked how larger/smaller is the risk in the missing participants compared with that of the observed participants for each treatment group and study. Giving a range of plausible value for the ratio or difference of risks may help quantify uncertainty around IMOR. The IMOR approach is implemented in STATA17 in the metanmis command18 (http://www.mrc-bsu.cam.ac.uk/software/stata-software/). The IMOR approach can incorporate the best-case/worst-case scenarios as special cases and has several advantages. It does not aim to estimate the missing outcomes but aims at making valid inference on the summary treatment effects.19 It also accounts for the uncertainty induced by missing outcome data, unlike the naive approaches that consider the imputed values as if they were fully observed. An obvious downside of this approach is its complexity compared with the naive approaches as it requires involvement of a knowledgeable statistician and expert opinion.

RESULTS
The approaches described above are illustrated via an example of a systematic review of all registration studies which compared the effectiveness of an atypical antipsychotic (amisulpride) with a typical one (haloperidol or flupenthixol) in schizophrenia.20–24 The condition of the participants was measured with the Brief Psychiatric Rating Scale (BPRS). The primary outcome is a 50% reduction in BPRS between baseline and week 6. We carried out a meta-analysis and results are reported here.

In figure 1 we present the ORs and the weights used in a random-effects meta-analysis reporting corresponding estimates for each one of the five studies, using complete cases meta-analysis (black
Möller 1997 study which has a large missing rate and an OR that model is attributable to the fact that IMOR reduces the weight of outcome data and consequently the MAR assumption cannot be evaluated. The random-effects model is assumed. The summary estimate does not necessarily increase when a method increases within-study variability, it may also result in significance between the complete cases analysis and the IMOR approach. The LOCF method provides the most precise estimate, while the Gamble and Hollis uncertainty interval is very wide and probably there is not enough power to detect a non-zero effect. The differences in significance between the complete cases analysis and the IMOR model is attributable to the fact that IMOR reduces the weight of Möller 1997 study which has a large missing rate and an OR that favours (though not significant) the typical antipsychotics. Although the IMOR approach increases within-study variability, it may also result in a decrease in between-study variability and uncertainty around the summary estimate does not necessarily increase when a random-effects model is assumed.

**DISCUSSION**

Meta-analysts typically do not have access to the reasons for missing outcome data and consequently the MAR assumption cannot be tested empirically. A sensitivity analysis is the only viable way to evaluate the effect of different scenarios for the missing data mechanism. Missing data should not be ignored and systematic reviewers should not take for granted that the problem has been appropriately handled in the trial level. Naïve imputation techniques such as imputing the mean value, LOCF and the best-case and worst-case scenarios, though widely used, may produce biased results and underestimated SEs. The uncertainty interval has low power.

Assuming a statistical model that relates treatment effect in missing data to those of the observed data is helpful to consider how robust results are to departures from the MAR assumption. Clinical expertise may inform such a model or a sensitivity analysis can be employed assuming increasingly more stringent scenarios.

The Cochrane Handbook urges systematic reviewers to contact the original investigators to request missing data; state explicitly the assumptions of any methods used to cope with missing values; perform sensitivity analysis to assess how robust results are; address the potential impact of missing data on the findings of the review.

In most trials, most of all in mental health trials, missing data are likely not to be missing at random. Researchers should document the reasons why data are missing and collect data on auxiliary variables that may be predictive of both the outcome and the probability of dropping out.

**Competing interests**

DM and GS received research funding from the European Research Council (IMMA 260559), AC and OE received funding from Greek national funds through the Operational Program ‘Education and Lifelong Learning’ of the National Strategic Reference Framework (NSRF)—Research Funding Program: ARISTEIA. Investing in knowledge society through the European Social Fund.

doi:10.1136/eb-2014-101900

**REFERENCES**


---

**Figure 2** Meta-analytical results for the ORs of short-term improvement in schizophrenia (response) for amisulpride versus other typical antipsychotics: various alternative approaches for handling missing outcome data.

