Missing data in trial-based cost-effectiveness analysis: An incomplete journey

Baptiste Leurent1 | Manuel Gomes2 | James R. Carpenter1,3

1 Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, UK
2 Department of Health Services Research and Policy, London School of Hygiene and Tropical Medicine, London, UK
3 MRC Clinical Trials Unit, University College London, London, UK

Correspondence
Baptiste Leurent, Department of Medical Statistics, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK.
Email: baptiste.leurent@lshtm.ac.uk

Summary
Cost-effectiveness analyses (CEA) conducted alongside randomised trials provide key evidence for informing healthcare decision making, but missing data pose substantive challenges. Recently, there have been a number of developments in methods and guidelines addressing missing data in trials. However, it is unclear whether these developments have permeated CEA practice. This paper critically reviews the extent of and methods used to address missing data in recently published trial-based CEA.

Issues of the Health Technology Assessment journal from 2013 to 2015 were searched. Fifty-two eligible studies were identified. Missing data were very common; the median proportion of trial participants with complete cost-effectiveness data was 63% (interquartile range: 47%–81%). The most common approach for the primary analysis was to restrict analysis to those with complete data (43%), followed by multiple imputation (30%). Half of the studies conducted some sort of sensitivity analyses, but only 2 (4%) considered possible departures from the missing-at-random assumption.

Further improvements are needed to address missing data in cost-effectiveness analyses conducted alongside randomised trials. These should focus on limiting the extent of missing data, choosing an appropriate method for the primary analysis that is valid under contextually plausible assumptions, and conducting sensitivity analyses to departures from the missing-at-random assumption.

Keywords
cost-effectiveness analysis, missing data, multiple imputation, randomised controlled trials, sensitivity analysis

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2018 The Authors Health Economics published by John Wiley & Sons Ltd.
1 | INTRODUCTION

Cost-effectiveness analyses (CEA) conducted alongside randomised controlled trials are an important source of information for health commissioners and decision makers. However, clinical trials rarely succeed in collecting all the intended information (Bell, Fiero, Horton, & Hsu, 2014), and inappropriate handling of the resulting missing data can lead to misleading inferences (Little et al., 2012). This issue is particularly pronounced in CEA because these usually rely on collecting rich, longitudinal information from participants, such as their use of healthcare services (e.g., Client Service Receipt Inventory; Beecham & Knapp, 2001) and their health-related quality of life (e.g., EQ-5D-3L; Brooks, 1996).

Several guidelines have been published in recent years on the issue of missing data in clinical trials (National Research Council, 2010; Committee for Medicinal Products for Human Use (CHMP), 2011; Burzykowski et al., 2010; Carpenter & Kenward, 2007) and for CEA in particular (Briggs, Clark, Wolstenholme, & Clarke, 2003; Burton, Billingham, & Bryan, 2007; Faria, Gomes, Epstein, & White, 2014; Manca & Palmer, 2005; Marshall, Billingham, & Bryan, 2009). Key recommendations include:

- taking practical steps to limit the number of missing observations;
- avoiding methods whose validity rests on contextually implausible assumptions, and using methods that incorporate all available information under reasonable assumptions; and
- assessing the sensitivity of the results to departures from these assumptions.

In particular, following Rubin’s taxonomy of missing data mechanisms (Little & Rubin, 2002), methods valid under a missing-at-random (MAR) assumption (i.e., when, given the observed data, missingness does not depend on the unseen values) appear more plausible than the more restrictive assumption of missing completely at random, where missingness is assumed to be entirely independent of the variables of interest. Because we cannot exclude the possibility that the missingness may depend on unobserved values (missing not at random [MNAR]), an assessment of the robustness of the conclusions to alternative missing data assumptions should also be undertaken.

Noble and colleagues (Noble, Hollingworth, & Tilling, 2012) have previously reviewed how missing resource use data were addressed in trial-based CEA. They found that practice fell markedly short of recommendations in several aspects. In particular, that reporting was usually poor and that complete-case analysis was the most common approach. However, missing data research is a rapidly evolving area, and several of the key guidelines were published after that review. Therefore, we aimed to review how missing cost-effectiveness data were addressed in recent trial-based CEA.

We reviewed studies published in the National Institute for Health Research Health Technology Assessment (HTA) journal, as it provides an ideal source for assessing whether recommendations have permeated CEA practice. These reports give substantially more information than a typical medical journal article, allowing authors the space to clearly describe the issues raised by missing data in their study and the methods they used to address these. Our primary objectives were to determine the extent of missing data, how these were addressed in the analysis, and whether sensitivity analyses to different missing data assumptions were performed. We also provide a critical review of our findings and recommendations to improve practice.

2 | METHODS

The PubMed database was used to identify all trial-based CEA published in HTA between the January 1, 2013, and December 31, 2015. We combined search terms such as “randomised,” “trial,” “cost,” or “economic” to capture relevant articles (see Appendix A.1 for details of the search strategy). The full reports of these articles were downloaded then screened for eligibility by excluding all studies that were pilot or feasibility studies; reported costs and effects separately (e.g., cost-consequence analysis); or did not report a within-trial CEA.

For each included study, we extracted key information about the study and the analysis to answer our primary research questions. A detailed definition of each indicator extracted is provided in Appendix B. In a second stage, we drew on published guidelines and our experience to derive a list of recommendations to address missing data, and then re-reviewed the studies to assess to which extent they followed these recommendations (see Appendix B for further details).

Data analysis was conducted with Stata version 15 (StataCorp, 2017). The data from this review are available on request (Leurent, Gomes, & Carpenter, 2017).
3 | RESULTS

3.1 | Included studies

Sixty-five articles were identified in our search (Figure 1), and 52 eligible studies were included in the review (listed in Appendix A.2). The median time frame for the CEA was over 12 months, and the majority of trials (71%, n = 37) conducted a follow-up with repeated assessments over time (median of 2; Table 1). The most common effectiveness measure was the quality-adjusted life year (81%, n = 42). Other outcomes included score on clinical measures, or dichotomous outcomes such as “smoking status”.

3.2 | Extent of missing data

Missing data was an issue in almost all studies, with only five studies (10%) having less than 5% of participants with missing data. The median proportion of complete cases was 63% (interquartile range, 47–81%; Figure 2). Missing data arose mostly from patient-reported (e.g., resource use and quality of life) questionnaires. The extent of missing data was generally similar for cost and effectiveness data, but 10 (19%) studies had more missing data in the latter (Table 1). The proportion of complete cases reduced, as the number of follow-up assessments increased (Spearman’s rank correlation coefficient $\rho = -0.59$, $p$ value < .001) and as the study duration increased ($\rho = -0.29$, $p = .04$).

3.3 | Approach to missing data

In the remaining assessments, we excluded the five studies with over 95% of complete cases. Three main approaches to missing data were used: complete-case analysis (CCA; Faria et al., 2014), reported in 66% of studies (n = 31), multiple imputation (MI; Rubin, 1987; 49%, n = 23), and ad hoc hybrid methods (17%, n = 8). For the primary analysis, CCA

---

**FIGURE 1** Studies selection flow diagram. CEA = cost-effectiveness analyses; HTA = health technology assessment; RCT = randomised controlled trial
was the most commonly used method (43%, n = 20), followed by MI (30%, n = 14; Table 2). MI was more common when the proportion of missing data was high and when there were multiple follow-up assessments (see Table 3).

### 3.4 Sensitivity analyses

Over half of the studies (53%, n = 25) did not conduct any sensitivity analysis around missing data, with 21% (n = 10) reporting CCA results alone and 11% (n = 5) MI results under MAR alone (Table 4). The remaining studies (n = 22, 47%) assessed the sensitivity of their primary analysis results to other approaches for the missing data. This was usually performing either MI under MAR, or CCA, when the other approach was used in the primary analysis. Other sensitivity analyses included using last observation carried forward or regression imputation.

Only two studies (4%) conducted sensitivity analyses, assuming data could be MNAR. In both studies, values imputed under a standard MI were modified to incorporate possible departures from the MAR assumption for both
FIGURE 2  Proportion of trial participants with complete data for the primary cost-effectiveness analysis. Shown for cost-effectiveness (n = 52), effectiveness (n = 47, unclear in 5 studies), and cost data (n = 40, unclear in 12 studies)

TABLE 2  Methods for handling missing data in primary analysis (n = 47)

<table>
<thead>
<tr>
<th>Primary analysis method</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete-case analysis</td>
<td>20</td>
<td>43</td>
</tr>
<tr>
<td>Multiple imputation</td>
<td>14</td>
<td>30</td>
</tr>
<tr>
<td>Other—single methods</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inverse probability weighting</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Bayesian model, missing data as unknown parameter</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Other—ad hoc hybrid methodsa</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>Using a combination of</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean imputationb</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Regression imputationc</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Inverse probability weightingd</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Assuming failure when outcome missing</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Multiple imputation</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Last observation carried forward</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Unclear</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>

aAd hoc hybrid method = several approaches to missing data combined, for example, using mean imputation for missing individual resource use items and multiple imputation for fully incomplete observations.

bMean imputation = replacing missing values by the average across other participants.

cRegression imputation = replace missing values by predicted value based on observed variables.

dInverse probability weighting = analysing complete data, weighted according to their modelled probability of being observed. These methods are presented in more details in other references (Baio & Leurent, 2016; Faria et al., 2014).
the cost and effectiveness data using a simplified pattern-mixture model approach (Faria et al., 2014; Leurent et al., 2018). The studies then discussed the plausibility of these departures from MAR and their implications for the cost-effectiveness inferences.

3.5 | Recommendations criteria

Table 5 reports the number of studies that reported evidence of following the recommendations from Figure 3 (see Section 4). Most studies reported being aware of the risk of missing data, for example, by taking active steps to reduce them (n = 35, 74%). In addition, almost two-thirds of the studies (n = 29, 62%) reported the breakdown of missing data by arm, time point, and endpoint. Only about one-third of the studies have clearly reported the reasons for the missing
Only one study (2%) appropriately discussed the implications of missing data in their cost-effectiveness conclusions.

### TABLE 5

Review of indicators based on recommendations criteria \((n = 47)\)

<table>
<thead>
<tr>
<th>Criterion(^a)</th>
<th>Met(^b)</th>
<th>Not met</th>
<th>Unclear</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n)</td>
<td>%</td>
<td>(n)</td>
</tr>
<tr>
<td><strong>Prevent</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1. Maximise response rate</td>
<td>35</td>
<td>74</td>
<td>12</td>
</tr>
<tr>
<td>A2. Alternative data sources</td>
<td>10</td>
<td>21</td>
<td>37</td>
</tr>
<tr>
<td>A3. Monitor completeness</td>
<td>17</td>
<td>36</td>
<td>30</td>
</tr>
<tr>
<td><strong>Primary</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B1. Assumption for primary analysis</td>
<td>17</td>
<td>36</td>
<td>27</td>
</tr>
<tr>
<td>B2. Appropriate primary method</td>
<td>17</td>
<td>36</td>
<td>27</td>
</tr>
<tr>
<td><strong>Sensitivity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1. Discuss departures from the primary assumption</td>
<td>0</td>
<td>0</td>
<td>47</td>
</tr>
<tr>
<td>C2. Consider broad range of assumptions</td>
<td>2</td>
<td>4</td>
<td>45</td>
</tr>
<tr>
<td>C3. Method valid under these assumptions</td>
<td>2</td>
<td>4</td>
<td>45</td>
</tr>
<tr>
<td><strong>Report</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1. Missing data by endpoint, arm, and time point</td>
<td>29</td>
<td>62</td>
<td>18</td>
</tr>
<tr>
<td>D2. Discuss reasons for missing data</td>
<td>16</td>
<td>34</td>
<td>31</td>
</tr>
<tr>
<td>D3. Describe methods used and assumptions</td>
<td>17</td>
<td>36</td>
<td>30</td>
</tr>
<tr>
<td>D4. Conclusions in light of missing data</td>
<td>1</td>
<td>2</td>
<td>46</td>
</tr>
</tbody>
</table>

\(^{a}\)See Figure 3 and Appendix B for definition of each criterion.

\(^{b}\)Report demonstrates evidence of having followed this recommendation. Not met if the recommendation was not followed or not mentioned. Unclear if some suggestions the criteria may have been met but information not clear enough. See Appendix B for detailed definitions and methodology used.

**FIGURE 3** Recommendations for improving handling of missing data in trial-based cost-effectiveness analysis.

References: 1, Little et al., 2012; 2, Noble et al., 2012; 3, Faria et al., 2014; and 4, Carpenter and Kenward 2007

Data \((n = 16, 34\%)\) and the approach used for handling the missing data and its underlying assumptions \((n = 17, 36\%)\). Only one study (2%) appropriately discussed the implications of missing data in their cost-effectiveness conclusions.

## 4 | DISCUSSION

### 4.1 | Summary of findings

Missing data remain ubiquitous in trial-based CEA. The median proportion of participants with complete cost-effectiveness data was only 63%. This reflects the typical challenges faced by CEA of randomised controlled trials, which often rely on patient questionnaires to collect key resource use and health outcome data. Despite best efforts to ensure...
completeness, a significant proportion of nonresponse is likely. This is consistent with other reviews, which also found no reduction of the extent of missing data in trials over time (Bell et al., 2014).

CCA remains the most commonly used approach for handling missing data in trial-based CEA, in contrast to recommendations. This approach makes the restrictive assumption that, given the variables in the analysis model, the distributions of the outcome data are the same, whether or not those outcome data are observed. This approach is also problematic because it can result in a loss in precision, as it discards participants who have partially complete data postrandomisation and who can provide important information to the analysis. Other unsatisfactory approaches based on unrealistic assumptions, such as last observation carried forward and single imputation, are also occasionally used.

MI (Rubin, 1987) assuming MAR has been widely recommended for CEA (Briggs et al., 2003; Burton et al., 2007; Faria et al., 2014; Marshall et al., 2009), allowing for baseline variables and postrandomisation data not in the primary analysis to be used for the imputation. It seems to be now more commonly used, with around half of the studies using MI for at least one of their analyses (up to 74% in 2015). Around one-third of the studies used MI for their primary CEA, which is higher than seen in primary clinical outcome analyses (8%; Bell et al., 2014).

On the other hand, sensitivity analyses to missing data remain clearly insufficient. Only two studies (4%) conducted comprehensive sensitivity analyses and assessed whether the study’s conclusions were sensitive to departures from the MAR assumption (i.e., possible MNAR mechanisms). Half of the studies did not conduct any sensitivity analysis regarding the missing data. The remaining studies performed some sort of sensitivity analyses, but usually consisting of simple variations from the primary analysis, such as reporting CCA results in addition to MI. This may be more for completeness than proper missing data sensitivity analyses. For example, if MI is used for the primary analysis (having assumed that MAR is the realistic primary missing data assumption), a sensitivity analysis that involves CCA will make stronger missing data assumptions.

### 4.2 Strengths and limitations

Our review follows naturally from the review of Noble et al. (2012) and gives an update of the state of play after the publication of several key guidelines. Our review, however, differs in scope and methods and cannot be directly compared with the results of Noble et al. One of the key strengths of this review is that HTA comprehensive reports allowed us to obtain a more complete picture of the missing data and the methods used to tackle it. HTA monographs are published alongside more succinct peer-reviewed papers in specialist medical journals, and they are often seen as the “gold-standard” for trial-based CEA in the UK. It seems therefore reasonable to assume that these are representative of typical practice in CEA. This review is, to our knowledge, the first to look at completeness of both cost and effectiveness data. A limitation is the use of a single-indicator “proportion of complete cases” to capture the extent of the missing data issue. This is however a clearly defined indicator and allows comparison with other reviews. The “recommendations indicators” also focused on the information reported in the study, not necessarily what might have been done in practice.

### 4.3 Recommendations

A list of recommendations to address missing data in trial-based CEA is presented in Figure 3. Trial-based CEA are prone to missing data, and it is important that analysts take active steps at the design and data-collection stages to limit their extent (Bernhard et al., 2006; Brueton et al., 2013; National Research Council, 2010). Resource use questionnaires should be designed in a user-friendly way, and their completion encouraged during follow-up visits, possibly supported by a researcher (Mercieca-Bebber et al., 2016; National Research Council, 2010). Alternative sources should also be considered to minimise missing information, for example, administrative data or electronic health records (Franklin & Thorn, 2018; Noble et al., 2012).

For any study with missing data, clear reporting of the issue is required. Ideally, the study should report details of the pattern of missing data (Faria et al., 2014), possibly as an appendix. At a minimum, CEA studies should report for each analysis the number of participants included by trial arm, as recommended in the Consolidated Standards of Reporting Trials guidelines (Noble et al., 2012; Schulz et al., 2010).

Although CCA may be justifiable in some circumstances, the choice of CCA for the primary analysis approach appears difficult to justify in the presence of repeated measurements, because the loss of power (by discarding all patients with any missing values) across the different time points tends to be large. Other approaches valid under more plausible
MAR assumptions and making use of all the observed data, such as MI (Rubin, 1987); likelihood-based repeated measures models (Faria et al., 2014; Verbeke, Fieuws, Molenberghs, & Davidian, 2014); or Bayesian models (Ades et al., 2006), should be considered. In particular, MI has been increasingly used in CEA, and further guidance to support an appropriate use in this context is warranted.

An area with clear room for improvement is the conduct of sensitivity analyses. This review found that many studies used CCA for the primary analysis and MI as a sensitivity analysis, or vice-versa, and concluded that the results were robust to missing data. This is misleading because both of these methods rely on the assumption that the missingness is independent of the unobserved data. Although the MAR assumption provides a sensible starting point, it is not possible to determine the true missing-data mechanism from the observed data. Studies should therefore assess whether their conclusions are sensitive to possible departures from that assumption (National Research Council, 2010; Committee for Medicinal Products for Human Use (CHMP), 2011; Faria et al., 2014). Several approaches have been suggested to conduct analyses under MNAR assumptions. Selection models express how the probability of being missing is related to the value itself. Pattern-mixture models, on the other hand, capture how missing data could differ from the observed (Molenberghs et al., 2014; Ratitch, O’Kelly, & Tosiello, 2013). Pattern-mixture models appear attractive because they frame the departure from MAR in a way that can be more readily understood by clinical experts and decision makers and can be used with standard analysis methods such as MI (Carpenter & Kenward, 2012; Ratitch et al., 2013). MNAR modelling can be challenging, but accessible approaches have also been proposed (Faria et al., 2014; Leurent et al., 2018). Further developments are still needed to use these methods in the CEA context and to provide the analytical tools and practical guidance to implement them in practice.

5 | CONCLUSION

Missing data can be an important source of bias and uncertainty, and it is imperative that this issue is appropriately recognised and addressed to help ensure that CEA studies provide sound evidence for healthcare decision making. Over the last decade, there have been some welcome improvements in handling missing data in trial-based CEA. In particular, more attention has been devoted to assessing the reasons for the missing data and adopting methods (e.g., MI) that can incorporate those in the analysis. However, there is substantial room for improvement. Firstly, more efforts are needed to reduce missing data. Secondly, the extent and patterns of missing data should be more clearly reported. Thirdly, the primary analysis should consider methods that make contextually plausible assumptions rather than resort automatically to CCA. Lastly, sensitivity analyses to assess the robustness of the study’s results to potential MNAR mechanisms should be conducted.

CONFLICT OF INTEREST

The authors have no conflict of interest.

ORCID

Baptiste Leurent http://orcid.org/0000-0001-6420-6567
Manuel Gomes http://orcid.org/0000-0002-1428-1232

REFERENCES


APPENDIX A
DETAILS OF STUDIES SELECTION

A.1 PubMed search criteria and results

Search criteria:

1. Search ("Health Technol Assess"[Journal])
   AND ("2013/01/01"[Date - Publication]: "2015/12/31"[Date - Publication])
   AND ("randomised"[Title OR "randomized"[Title] OR "trial"[Title]])
   AND ("economic"[Title/Abstract] OR "cost"[Title/Abstract])
   NOT ("pilot"[Title] OR "feasibility"[Title])

Number of studies:

<table>
<thead>
<tr>
<th>Search Query</th>
<th>Items found</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. Search (&quot;Health Technol Assess&quot;[Journal]) AND (&quot;2013/01/01&quot;[Date - Publication]: &quot;2015/12/31&quot;[Date - Publication]) AND (&quot;randomised&quot;[title] OR &quot;randomized&quot;[Title] OR &quot;trial&quot;[Title]) AND (&quot;economic&quot;[Title/Abstract] OR &quot;cost&quot;[Title/Abstract])</td>
<td>74</td>
</tr>
<tr>
<td>2. Search (&quot;Health Technol Assess&quot;[Journal]) AND (&quot;2013/01/01&quot;[Date - Publication]: &quot;2015/12/31&quot;[Date - Publication]) AND (&quot;randomised&quot;[Title] OR &quot;randomized&quot;[Title] OR &quot;trial&quot;[Title])</td>
<td>91</td>
</tr>
<tr>
<td>1. Search (&quot;Health Technol Assess&quot;[Journal]) AND (&quot;2013/01/01&quot;[Date - Publication]: &quot;2015/12/31&quot;[Date - Publication])</td>
<td>236</td>
</tr>
</tbody>
</table>

A.2 Included studies


APPENDIX B
INDICATORS DEFINITION

B.1 Primary indicators

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Definition</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of complete cases</td>
<td>Proportion of randomised participants for whom all data were available for the primary cost-effectiveness analysis</td>
<td>If the number of complete-cases was not clearly reported, we estimated an “upper bound,” from information, such as the proportion of participants with complete cost, or effect, data. See definition of primary analysis below.</td>
</tr>
<tr>
<td>Proportion complete effectiveness data</td>
<td>Proportion of randomised participants for whom all effectiveness data were available for the primary cost-effectiveness analysis</td>
<td>Same as above</td>
</tr>
<tr>
<td>Proportion complete cost data</td>
<td>Proportion of randomised participants for whom all cost data were available for the primary cost-effectiveness analysis</td>
<td>Same as above</td>
</tr>
<tr>
<td>Report exact number of complete cases</td>
<td>Whether the number of participants with complete cost and effectiveness data was clearly reported.</td>
<td></td>
</tr>
<tr>
<td>More missing costs or effectiveness</td>
<td>Whether the proportion of complete cases differ between cost and effectiveness variable.</td>
<td>Considered “similar” when the proportion of complete cases was within 5% of each other.</td>
</tr>
<tr>
<td>Primary analysis method</td>
<td>Methods used to address missing data in the primary (base case) cost-effectiveness analysis</td>
<td>When multiple effectiveness measures, time-frames, or cost perspectives were reported, without a base-case clearly defined, we considered the analysis based on quality-adjusted life years (QALYs) over the longest within-trial follow-up period, from the NHS and social services cost perspective.</td>
</tr>
<tr>
<td>Conducted a sensitivity analysis to missing data</td>
<td>Report results under more than one approach for addressing missing data</td>
<td></td>
</tr>
</tbody>
</table>
B.2 | Secondary indicators: Derived from the recommendations list

B.2.1 | Methods

Because these aspects could have been mentioned in multiple parts in the monograph, we used a systematic approach, looking for keywords and checking the most relevant paragraphs in the full report.

i. Search in PDF: “Missing”; “Participation”; “Completion”; “Incomplete”; “Response”; “Non-response”; “Monitor”; “MCAR”; “MAR”; “MNAR.”

ii. If did not find “steps to reduce missing data,” also check “reminder,” “incentive,” “telephone,” and “contact.”

iii. Then, check relevant paragraphs manually: data source for cost-effectiveness data; beginning of CEA results; and CEA conclusions.

B.2.2 | Answers

“Yes”: The recommendation was clearly mentioned, and the criteria therefore met.

“No”: The recommendation was not clearly mentioned or found. The recommendation may still have been followed but not reported (or at least not found with the above strategy).

“Unclear”: There was some suggestions the criteria may have been met but not enough information to be sure.

<table>
<thead>
<tr>
<th>Recommendation</th>
<th>Indicator definition</th>
<th>Examples “yes”</th>
<th>Examples “no”</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1. Maximise response rate (consider questionnaire design, mode of administration, reminders, incentives, participants engagement, etc.)</td>
<td>Mention taking steps to maximise response rate</td>
<td>Reminder, incentives, home/hospital visit, multiple attempts,</td>
<td>Mention response was maximised for clinical outcome but not reported for cost-effectiveness endpoints</td>
<td>Can be for overall trial data if implicit includes cost or effect data. Except if steps are clearly for non-CE variables only (e.g., primary outcome only).</td>
</tr>
<tr>
<td>A2. Consider alternative data sources (e.g., routinely collected data)</td>
<td>Mention that considered missing data issues when choosing appropriate source, OR mention more than one source used for a CE data.</td>
<td>Use of electronic health records or administrative data, e.g., hospital episode statistics were used to supplement trial’s data, for example, about hospital admissions post-randomisation (which might be otherwise missing).</td>
<td>Using routine data as a primary source: e.g., resource use taken primarily from administrative/hospital records.</td>
<td></td>
</tr>
<tr>
<td>A3. Monitor cost-effectiveness data completeness while trial ongoing</td>
<td>Mentioned monitoring data completeness while trial ongoing.</td>
<td>Data managers checked inconsistent and missing data (if not clear “while trial ongoing” but mention monitoring probably fine). Mention taking new steps to reduce MD (e.g., incentive) as realised lots of MD after trial started.</td>
<td>Mention data checks for inconsistencies, but no mention of checking missing data.</td>
<td>Can be for overall trial data. Except if monitoring clearly for non-CE variables only (e.g., primary outcome only).</td>
</tr>
<tr>
<td>B1. Formulate realistic and accessible missing data assumption for the primary analysis (typically, but not necessarily, a form of the missing at random assumption)</td>
<td>Primary (base-case) CEA based on reasonable missing data assumptions. (likely MAR, or alternative if well justified).</td>
<td>– Used MI for primary analysis - well justified and clear alternative</td>
<td>– Hybrid method, except if clearly explain and justify underlying assumptions</td>
<td></td>
</tr>
</tbody>
</table>

(Continues)
<table>
<thead>
<tr>
<th>Recommendation</th>
<th>Indicator definition</th>
<th>Examples “yes”</th>
<th>Examples “no”</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>B2. Use appropriate method valid under that assumption (typically, but not necessarily, multiple imputation or maximum likelihood)</td>
<td>Use appropriate analysis method.</td>
<td>– MI for primary analysis - Bayesian under MAR - well justified and clear alternative</td>
<td>– Use unadjusted CCA when reporting data are MAR.</td>
<td></td>
</tr>
<tr>
<td>C1. Discuss with clinicians and investigators to formulate plausible departures from the primary missing data assumption</td>
<td>Conducted MNAR SA + mention elicitation.</td>
<td>Did not conduct MNAR SA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2. Consider a broad range of assumptions, including missing not at random mechanisms</td>
<td>Conducted MNAR SA</td>
<td>Did not conduct MNAR SA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3. Use appropriate method valid under these assumptions (typically, but not necessarily, pattern-mixture models or reference-based approach)</td>
<td>Conducted MNAR SA, and used an appropriate method (PMM, etc.).</td>
<td>Did not conduct MNAR SA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D1. Report number of participants with cost and outcome data, by arm and time-point</td>
<td>Report number (or %) of complete or missing data. Split at least by effectiveness vs. cost, time point (when applicable), and arm</td>
<td>Reported missing data by endpoint and arm, but not by time point.</td>
<td>Do not have to be all at the same time (split by endpoint + time + arm), can be three separate table/texts.</td>
<td></td>
</tr>
<tr>
<td>D2. Report possible reasons for non-response, and baseline predictors of missing values</td>
<td>Mention something about main reason for the missing data, OR Explore factors associated with it.</td>
<td>Comment on why missing data (e.g., “because patients were too ill”). Or explore baseline factors associated with missingness</td>
<td>No mention of reasons for MD in the CE section.</td>
<td>Have to be specific to the CE missing data, or clearly mentioning something like &quot;reasons for MD are discussed in clinical analysis section ...&quot;</td>
</tr>
<tr>
<td>D3. Describe methods used, and underlying missing data assumptions</td>
<td>Clearly state the method used to address missing data, AND the underlying assumption.</td>
<td>No report of missing data assumption or method used</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Draw overall conclusion in light of the different results and the plausibility of the respective assumptions</td>
<td>Conduct sensitivity analyses, and interpret results appropriately.</td>
<td>Did MNAR SA and appropriate conclusion.</td>
<td>– Did not conduct sensitivity analyses</td>
<td>– Did not conduct sensitivity analyses, but no comment/conclusion</td>
</tr>
</tbody>
</table>