Journal of Clinical Epidemiology 99 (2018) 14–23

REVIEW

Systematic reviews do not adequately report or address missing outcome data in their analyses: a methodological survey

Lara A. Kahale[a], Batoul Diab[a], Romina Brignardello-Petersen[b,c], Arnav Agarwal[b,d], Reem A. Mustafa[b,e,f], Joey Kwong[g], Ignacio Neumann[b,h], Ling Li[i], Luciane Cruz Lopes[j], Matthias Briel[b,k], Jason W. Busse[b,l,m], Alfonso Iorio[b,n], Per Olav Vandvik[o,p], Paul Elias Alexander[b], Gordon Guyatt[b,n], Elie A. Akl[a,b,*]

[a]*Clinical Epidemiology Unit, American University of Beirut, Beirut, Lebanon*
[b]*Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, Ontario, Canada*
[c]*Evidence Based Dentistry Unit, Faculty of Dentistry, Universidad de Chile, Santiago, Chile*
[d]*Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada*
[e]*Department of Medicine, University of Missouri-Kansas City, Kansas City, MO, USA*
[f]*Department of Biomedical and Health Informatics, University of Missouri-Kansas City, Kansas City, MO, USA*
[g]*JC School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong, China*
[h]*Department of Internal Medicine, School of Medicine, Pontificia Universidad Católica de Chile, Santiago, Chile*
[i]*Clinical Research and Evaluation Unit, Chinese Evidence-based Medicine Center, West China Hospital, Sichuan University, Chengdu, China*
[j]*Pharmaceutical Sciences Post Graduate Course, University of Sorocaba, UNISO, Sorocaba, Sao Paulo, Brazil*
[k]*Basel Institute for Clinical Epidemiology and Biostatistics, Department of Clinical Research, University Hospital Basel, University of Basel, Spitalstrasse 12 4031, Basel, Switzerland*
[l]*Department of Anesthesia, McMaster University, Hamilton, Canada*
[m]*The Michael G. DeGroote Institute for Pain Research and Care, McMaster University, Hamilton, Canada*
[n]*Department of Medicine, McMaster University, Hamilton, Canada*
[o]*Norwegian Knowledge Centre for the Health Services, Oslo, Norway*
[p]*Department of Medicine, Innlandet Hospital Trust, Gjøvik, Norway*

Accepted 24 February 2018; Published online 2 March 2018

Abstract

Objectives: To describe how systematic review authors report and address categories of participants with potential missing outcome data of trial participants.

Study Design and Setting: Methodological survey of systematic reviews reporting a group-level meta-analysis.

Results: We included a random sample of 50 Cochrane and 50 non-Cochrane systematic reviews. Of these, 25 reported in their methods section a plan to consider at least one of the 10 categories of missing outcome data; 42 reported in their results, data for at least one category of missing data. The most reported category in the methods and results sections was ''unexplained loss to follow-up'' ($n = 34$ in methods section and $n = 6$ in the results section). Only 19 reported a method to handle missing data in their primary analyses, which was most often complete case analysis. Few reviews ($n = 9$) reported in the methods section conducting sensitivity analysis to judge risk of bias associated with missing outcome data at the level of the meta-analysis; and only five of them presented the results of these analyses in the results section.

Conclusion: Most systematic reviews do not explicitly report sufficient information on categories of trial participants with potential missing outcome data or address missing data in their primary analyses.

Keywords: Missing outcome data; Imputation; Risk of bias; Trials; Systematic reviews; Meta-analysis

Conflicts of interest statement: The authors report that they have no conflicts of interest to disclose.

Funding: The study was funded by Cochrane Methods Innovation Fund.

Authors' contributions: L.A.K., G.G., and E.A.A. contributed to the conception and design of the manuscript. P.E.A. contributed to study selection. L.A.K. and B.D. contributed to the full text screening. L.A.K., B.D., R.B.P., A.A., R.A.M., J.K., I.N., L.L., L.C.L., M.B., J.W.B., A.I., and P.O.V. contributed to data abstraction. L.A.K. and E.A.A. contributed to data analysis, interpretation of results, and drafting of the manuscript. All authors revised and approved the manuscript.

* Corresponding author. Department of Internal Medicine, Clinical Epidemiology Unit, American University of Beirut Medical Center, P.O. Box: 11-0236, Riad-El-Solh Beirut 1107 2020, Beirut, Lebanon. Tel.: +96176681010.

E-mail address: ea32@aub.edu.lb (E.A. Akl).

https://doi.org/10.1016/j.jclinepi.2018.02.016
0895-4356/

What's new?

Key findings

- Systematic review authors do not explicitly report sufficient information on categories of trial participants with potential missing outcome data.
- Systematic review authors do not typically handle missing data in their primary analyses.
- Systematic review authors do not typically judge risk of bias (ROB) associated with missing outcome data at the level of the meta-analysis.

What this adds to what is known?

- The most reported category among systematic review authors is "unexplained loss to follow-up".
- Complete case analysis is the most frequent method used by systematic review authors to handle missing data in their primary meta-analyses.

What is the implication, what should change now?

- Systematic review authors, journal reviewers, and editors need to better adhere to the PRISMA statement's recommendations related to reporting of missing outcome data.
- Systematic reviewers should assess the risk of bias associated with missing outcome data (MOD) at the level of the meta-analysis and test whether significance of effect estimates is or is not robust to MOD.

1. Introduction

Attrition bias is a frequent problem in the conduct of randomized trials. It refers to the potential bias introduced by participants who have missing outcome data for outcomes of interest. Eighty-seven percent of trials published in five top medical journals suffer from missing outcome data (MOD) [1]. Up to a third of positive trials in these prestigious journals lose statistical significance when one makes plausible assumptions about the outcomes of participants with MOD [1]. This bias is expected to affect the validity of findings not only of these trials but also of systematic reviews including them.

One approach for handling MOD in systematic reviews is to calculate a single credible estimate of treatment effect, together with an estimate of its uncertainty accounting for "the strength of evidence" and MOD [2]. This approach depends on the classification of MOD according to the relationship between missingness and observed or unobserved factors (e.g., missing completely at random, missing at random, and missing not at random) [3]. For the primary meta-analysis, the grading of recommendations assessment, development, and evaluation (GRADE) working group recommends either conducting a complete case analysis or making assumptions about the outcomes of those with MOD if investigators have strong hypotheses for those outcomes [4]. GRADE further recommends conducting sensitivity analyses using plausible assumptions for those with MOD to evaluate the robustness of the primary meta-analyses (i.e., assess the risk of bias) [4,5]. Despite the various suggested approaches, only a quarter of systematic reviews report a plan for handling MOD [6].

To apply optimal approaches to addressing MOD, systematic reviews need to identify which participants or categories of participants have MOD. This requires trialists to report whether participants belonging to categories of MOD (e.g., those who withdrew consent or discontinued treatment drug) were followed. In addition, primary study authors often do not make clear statements about their assumptions regarding MOD. For example, participants with MOD may have been excluded from the numerator and denominator (i.e., complete case analysis) or included in the denominator with assumptions made about their outcomes in the numerator (imputation).

These challenges may contribute to the apparently poor performance of systematic reviews in addressing MOD: a survey of systematic reviews published in the Cochrane Database of Systematic Reviews between 2009 and 2012 by three Cochrane Review Groups relating to mental health found that only 3% provided a clear definition of MOD [7]. The investigators recommended that systematic review authors, journal reviewers, and editors should ensure explicit definition of terms used to categorize participants with potentially MOD [7].

Given these apparent problems in systematic reviews, we conducted a systematic survey of reviews to further explore their performance with respect to MOD.

1.1. Objectives

The main objective of this study is to describe how systematic review authors report categories of participants with potential MOD. In addition, we assessed how authors of systematic reviews handle MOD in their primary meta-analyses of dichotomous outcomes and assess the associated risk of bias.

2. Methods

2.1. Design overview

This study is part of a larger project examining the reporting, handling, and assessment of risk of bias associated with MOD in trials and systematic reviews. We have reported details of the project's definitions and methodology elsewhere [8]. We used standard systematic review

methodology to conduct a survey of systematic reviews reports. Because this study involves no human participants, we have not sought ethical approval.

2.2. Eligibility criteria

An eligible systematic review met the following criteria:

- Is described as "systematic review" and/or "meta-analysis" of trials;
- Compares one clinical intervention to another (or to no intervention);
- Reports a search strategy of at least one electronic database;
- Addresses a preventive or a therapeutic clinical question in humans (diagnostic, prognostic, public health, and health services questions were not eligible);
- Is published in the Cochrane Database of Systematic Reviews or in a core clinical journal indexed in MEDLINE;
- Includes a meta-analysis meeting the following criteria:
    - Is a group level frequentist meta-analysis of randomized controlled trials and/or controlled clinical trials (e.g., network meta-analysis, Bayesian meta-analysis, and meta-regression alone are not eligible);
    - Reports an effect estimate expressed as a dichotomous measure (including relative risk or odds ratio with arm-level data;
    - Reports a statistically significant pooled effect estimate from at least two trials for a patient-important efficacy outcome; statistical significance refers to P-value <0.05 or confidence interval not including 1.0.

The exclusion criteria were as follows:

1. Duplicate publications (e.g., a Cochrane systematic review published in both the Cochrane Library and in a peer review journal);
2. A meta-analysis with more than 20 included trials to ensure feasibility;
3. A meta-analysis not reporting the numerator and denominator in each arm of each eligible study.

2.3. Search strategy

We searched the Cochrane Library for Cochrane systematic reviews and used the OVID Medline interface to search for non-Cochrane systematic reviews in the Core Clinical Journals (119 English language clinical journals indexed under Abridged Index Medicus by the National Library of Medicine; available at http://www.nlm.nih.gov/bsd/aim.html). The search included studies published in 2012, in any language. Appendix on the journal's web site at www.elsevier.com provides the details of our search strategy.

2.4. Selection process

The search strategy captured a total of 1,137 citations. We proceeded by screening successive random samples of 100 Cochrane reviews and 100 non-Cochrane reviews until we reached our desired sample size of 100 reviews (50 Cochrane and 50 non-Cochrane). We used an online tool (https://www.random.org/sequences/) to generate the random sequences that we used to create the random samples.

Teams of two reviewers, screened independently and in duplicate, titles and abstracts, and then full texts for eligibility. We conducted calibration exercises and used standardized and pilot-tested forms with detailed written instructions. For each review, we selected the first reported meta-analysis of the first reported patient-important efficacy outcome with significant pooled effect estimate referred to thereafter as selected meta-analyses.

2.5. Data abstraction

Five pairs of reviewers conducted data abstraction independently and in duplicate using web-based systematic review software (DistillerSR). We used a pilot-tested standardized data abstraction form and conducted calibration exercises. As planned in the protocol [8], we collected from each eligible systematic review, information relevant to the following: (1) characteristics of the systematic review, (2) reporting of MOD, (3) handling of MOD, and (4) assessment of risk of bias associated with MOD. We abstracted information about MOD in reference to the selected meta-analysis analysis of interest (i.e., comparison and outcome addressed in the selected meta-analysis) [8].

Regarding categories of participants with potential MOD, verified whether the systematic review authors:

- Explicitly planned as part of the methods section to consider the following 10 categories of potential MOD: (1) "ineligible participants or mistakenly randomized", (2) "did not receive first intervention", (3) "withdrew consent", (4) "explained lost to follow-up (LTFU)" (i.e., moved out of country), (5) "unexplained LTFU" (i.e., for reasons not considered in our other categories), (6) "noncompliant", (7) "discontinued trial prematurely", (8) "crossover", (9) "dead", (10) "adverse events", or other reasons. We initially referred to a previous list of potential MOD categories published elsewhere [1]; however, this list was continuously modified as we abstracted data depending on what was frequently reported among eligible reviews.
- Explicitly reported data for the aforementioned categories in the results section and at what level (e.g., at the study arm level, study level, and across studies).

Regarding handling of categories of participants with potential MOD, we verified whether the systematic review:

- Explicitly stated using specific analytical method(s) for addressing MOD in the primary analysis of the

selected meta-analysis (i.e., to account for MOD when generating the best effect estimate). These methods include the following: (1) complete case analysis, (2) making assumptions for MOD, (3) using the assumption(s) made by the trialists, or (4) excluding trials with high rates of MOD;
- Provided justification for the analytical method(s) used to handle MOD in the primary analysis of the selected meta-analysis.

Regarding assessing the risk of bias associated with MOD, we assessed whether the systematic review:

- Evaluated the risk of bias associated with MOD at the trial level and the tool used (e.g., Cochrane Risk of Bias [RoB] tool);
- Stated method(s) used to assess or judge risk of bias associated with MOD at the level of the meta-analysis, for example, sensitivity analysis and subgroup analysis;
- Provided the results of any sensitivity meta-analyses applied to account for MOD
- "Took into account the uncertainty associated with imputing events in the primary or secondary analysis". Imputing events require including participants with MOD in the denominator and making assumptions about their outcomes in the numerator. This naïve approach considers the imputed values as if they were fully observed, leading to a false narrowing of the confidence interval. To correct for this, methodologists have developed methods that take into account uncertainty associated with imputing missing observations using sophisticated statistical approaches [9–12].

2.6. Data analysis

We used the kappa statistic to calculate agreement between reviewers for the inclusion of systematic reviews at the full-text screening stage. We judged the level of agreement according to the guidelines proposed by Landis and Koch [13]: kappa values of 0–0.20 represent slight agreement, 0.21–0.40 fair agreement, 0.41–0.60 moderate agreement, 0.61–0.80 substantial agreement, and greater than 0.80 almost perfect agreement.

We conducted descriptive analysis for all variables; overall and stratified by Cochrane and non-Cochrane reviews. For categorical variables, we reported frequencies and percentages. For continuous variables, we used mean and standard deviation when data were normally distributed. Otherwise, we reported median and interquartile range. We used the Shapiro–Wilk test to evaluate whether distributions of continuous variables violated assumptions of normality.

We tested for the statistical significance of the differences between the Cochrane and non-Cochrane reviews for all relevant analyses. For dichotomous variables, we used the chi-square test, or Fisher's exact test if the expected event number was <5. For continuous variables, we used the Student's *t*-test for two independent samples when the distribution was normal, and the Mann-Whitney *U*-test, when the distribution was not normal. For all analyses, we used the SPSS statistical software, version 21.0 (SPSS INC, Chicago, Illinois, USA).

3. Results

Fig. 1 shows the study flow. Our electronic search identified a pool of 1,137 citations. From these, we selected 50 Cochrane and 50 non-Cochrane based on our eligibility criteria (refer to Appendix 1 for the list of journals). Agreement between authors for study eligibility was almost perfect (kappa = 0.8).

3.1. General characteristics of selected meta-analyses

Table 1 presents the characteristics of all included studies. Compared with non-Cochrane reviews, Cochrane reviews included fewer trials and less frequently addressed active pharmacological controls. Cochrane reviews more frequently reported the following: using the GRADE approach for rating certainty in estimates; the conduct of intention-to-treat analysis; and funding by government and private not-for-profit institutions. There was no significant difference in the rates of MOD between Cochrane and non-Cochrane reviews.

3.2. Reporting of categories of participants with potential MOD

Table 2 summarizes the reporting of information regarding categories of participants that could be potentially considered MOD. Of the Cochrane reviews, 44 reported an explicit plan in their methods section to consider at least one of these categories of MOD; six of the non-Cochrane reviews did so (*P*-value <0.001). Only 11 reviews explicitly planned to consider any MOD categories other than "unexplained LTFU", "explained LTFU", and other category.

Of the 100 reviews, 42 reported at least one of the MOD categories of interest in their results section (29 Cochrane and 13 non-Cochrane reviews). When provided, the number of participants with potential MOD was reported per trial and per arm in 45% of the Cochrane reviews and 15% of non-Cochrane reviews.

3.3. Handling of MOD

Table 3 shows how systematic reviewer authors handled MOD. Nineteen reviews reported a plan for handling MOD of dichotomous outcomes. The two most frequently reported approaches were complete case analysis and assuming no participants with MOD had the event (five and four reviews, respectively). Only one Cochrane review and one non-Cochrane systematic review provided justification for any of the methods used for handling MOD.

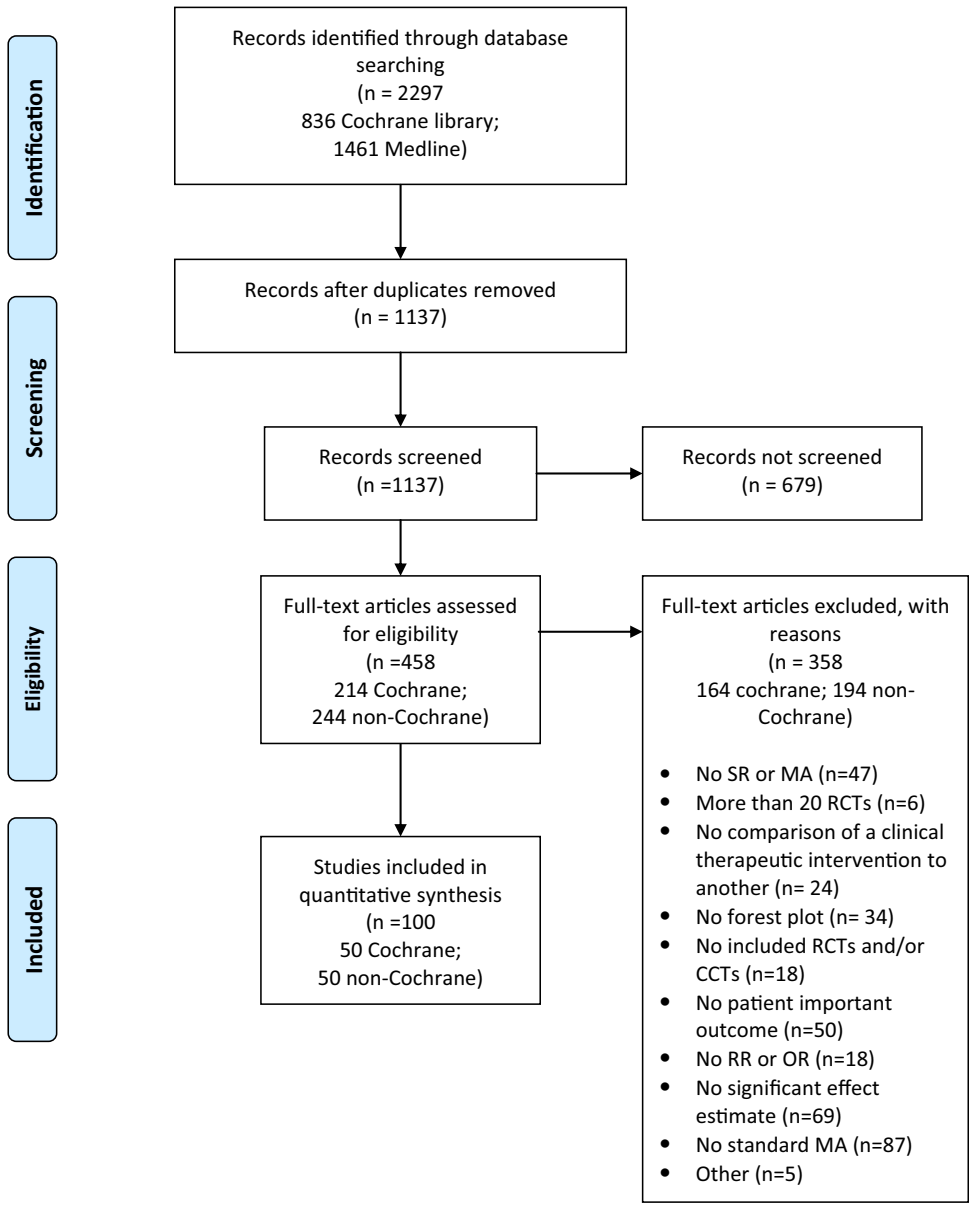

Fig. 1. PRISMA study flow diagram. OR, odds ratio; RR, relative risk. CCT: clinically controlled trial; MA: meta-analysis; SR: systematic review. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. https://doi.org/10.1371/journal.pmed1000097. For more information, visit http://www.consort-statement.org.

None of the systematic reviews that reported assumptions for addressing MOD took uncertainty into account.

3.4. Assessing risk of bias associated with MOD

Table 4 describes the assessment of the risk of bias associated with MOD. Of the nine reviews that reported in the Section 2 conducting a sensitivity analysis to assess the risk of bias associated with MOD at the level of the meta-analysis, only five presented the results of their analysis. One out of three Cochrane reviews which reported to conduct a second sensitivity analysis to judge risk of bias associated with MOD actually presented the results of the sensitivity analysis. Risk of bias associated with MOD at the level of the trial was assessed in 87 reviews: 65% used the Cochrane RoB tool, and 22% used a tool other than the Cochrane RoB tool. Out of the 65 reviews that used the Cochrane RoB, 86% were Cochrane, and 44% were non-Cochrane ($P < 0.001$).

4. Discussion

4.1. Summary of findings

Although 42 of 100 systematic reviews reported on at least one of the 10 predefined categories of MOD, only 19 reported plans for handling MOD in their analyses. Although 87 reviews addressed risk of bias associated with

Table 1. General characteristics of included systematic reviews

	Overall ($N = 100$)	Cochrane SR ($N = 50$)	Non-Cochrane SR ($N = 50$)	*P* value [a]
Number of trials included; median (IQR)	6 (3-8)	5 (3-7)	6.5 (4-9)	0.036[a]
Outcome category				
Mortality	21 (21%)	9 (18%)	12 (24%)	0.101
Morbidity	56 (56%)	25 (50%)	31 (62%)	
Patient reported outcomes	23 (23%)	16 (32%)	7 (14%)	
Type of Intervention				
Pharmacological	61 (61%)	33 (66%)	27 (56%)	0.473
Surgery /invasive procedure	24 (24%)	9 (18%)	15 (30%)	
Other	15 (15%)	8 (16%)	7 (14%)	
Type of control				
Active: pharmacological	21 (21%)	6 (12%)	15 (30%)	0.034[a]
Active: surgery /invasive procedure	18 (18%)	7 (14%)	11 (22%)	
Non-active: no intervention/ standard of care /placebo /sham	55 (55%)	35 (70%)	20 (40%)	
Other	6 (6%)	2 (4%)	4 (8%)	
Used the GRADE approach	29 (29%)	22 (44%)	7 (14%)	0.001[a]
Funding				
Private for profit	4 (4%)	2 (4%)	2 (4%)	0.691
Private not for profit	40 (40%)	31 (62%)	9 (18%)	<0.001[a]
Government	34 (34%)	21 (46%)	11 (22%)	0.011[a]
Not funded	16 (16%)	6 (12%)	10 (20%)	0.275
Not reported	23 (23%)	3 (6%)	20 (40%)	<0.001[a]
Duration of follow-up in months; mean (SD)	12.4 (23.1)	13 (31.1)	11.8 (11.7)	0.840
Explicitly stating using the following in the meta-analyses				
Analyze as randomized	3 (3%)	3 (6%)	0 (0%)	0.121
Intention-to-treat	31 (31%)	20 (40%)	11 (22%)	0.052[a]
Modified Intention-to-treat	0 (0%)	0 (0%)	0 (0%)	N/A
Per-protocol	4 (4%)	1 (2%)	3 (6%)	0.309
As treated	0 (0%)	0 (0%)	0 (0%)	N/A
None of the above reported	67 (67%)	30 (60%)	37 (74%)	0.137
Rate of MOD; median (IQR)	8 (3-17)	6.5 (1-15.5)	13 (3-21)	0.325

Abbreviations: IQR, interquartile range; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; MOD, missing outcome data; SD, standard deviation.

[a] *P*-value for the difference between Cochrane and non-Cochrane systematic reviews.

MOD at the trial level, only nine reported in the Section 2 conducting sensitivity analysis as a way to judge risk of bias associated with MOD at the level of the meta-analysis. Of these, only five reported the results of their analysis.

4.2. Strengths and limitations

The main strength of our study is the systematic and transparent methods used in conducting our methodological survey, including but not limited to screening in duplicate and independently, calibration exercises. Furthermore, we explored issues that have not previously been addressed, particularly in terms of categorizing MOD. Upto our knowledge, this is only the second methodological study on MOD in systematic reviews that explores the categories of participants that constitute participants with MOD [6]. Finally, we included Cochrane and non-Cochrane systematic reviews to make our results more generalizable and to explore possible differences.

One limitation of our study is the restriction of our search strategy to MEDLINE for the identification of non-Cochrane systematic reviews. However, these reviews represent those typically accessed by clinicians. We also exclusively focused on dichotomous outcomes given the methods for addressing MOD for continuous outcomes are less well developed [14]. Although we focused only on meta-analyses that included 20 or fewer trials for feasibility issues, we doubt that the findings would be different for larger meta-analyses with greater than 20 trials.

Another limitation of our study is that instead of conducting a formal sample size calculation and including

Table 2. Reporting of information regarding categories of participants that could be potentially counted to have missing outcome data in Cochrane and non-Cochrane systematic reviews

	Overall ($N = 100$)	Cochrane SR ($N = 50$)	Non-Cochrane SR ($N = 50$)	*P* value [a]
Explicitly planned as part of the methods section to consider the following categories as having MOD				
Ineligible participants/ mistakenly randomized	0 (0%)	0 (0%)	0 (0%)	-
Did not receive first intervention	0 (0%)	0 (0%)	0 (0%)	-
Withdrew consent	1 (1%)	1 (2%)	0 (0%)	0.5
Explained LTFU	6 (6%)	6 (12%)	0 (0%)	0.13
Unexplained LTFU	21 (21%)	18 (36%)	3 (6%)	<0.001[a]
Non-compliant	1 (1%)	1 (2%)	0 (0%)	0.5
Discontinued trial prematurely	0 (0%)	0 (0%)	0 (0%)	-
Cross-over	0 (0%)	0 (0%)	0 (0%)	-
Dead	1 (1%)	1 (2%)	0 (0%)	0.5
Adverse events	1 (1%)	0 (0%)	1 (2%)	0.5
Other	7 (7%)	5 (10%)	2 (4%)	0.22
None of the above	75 (75%)	28 (56%)	47 (94%)	<0.001[a]
Explicitly reported in results section data for the following categories with potential MOD				
Ineligible participants/ mistakenly randomized	6 (6%)	6 (12%)	0 (0%)	0.013[a]
Did not receive intervention	5 (5%)	4 (8%)	1 (2%)	0.18
Withdrew consent	9(9%)	8 (16%)	1 (2%)	0.015[a]
Explained LTFU	5 (5%)	5 (10%)	0 (0%)	0.03[a]
Unexplained LTFU	37 (37%)	25 (50%)	12 (24%)	0.007[a]
Non-compliant	8 (8%)	7 (14%)	1 (2%)	0.03[a]
Discontinued prematurely	10 (10%)	5 (10%)	5 (10%)	1.0
Cross-over	1 (1%)	1 (2%)	0 (0%)	0.5
Dead	6 (6%)	3 (6%)	3 (6%)	0.66
Adverse events	15 (15%)	10 (20%)	5 (10%)	0.16
Other	9 (9%)	6 (12%)	3 (6%)	0.24
None of the above	58 (58%)	21 (42%)	37(74%)	0.001[a]
Reported number of participants with MOD[b]				
For each trial, per arm	15 (36%)	13 (45%)	2 (15%)	0.001[a]
For each trial, overall (arms combined)	12 (29%)	8 (28%)	4 (31%)	
Across trials, per arm	2 (4.8%)	1 (3%)	1 (8%)	
Across trials, overall (arms combined)	4 (9.5%)	3 (10%)	1 (8%)	
No	9 (21%)	4 (14%)	5 (38%)	

Abbreviations: MOD, missing outcome data; LTFU, lost to follow-up; SR, systematic review.

[a] *P*-value for the difference between Cochrane and non-Cochrane systematic reviews.

[b] $N = 44$ trials that explicitly provided in results section data for any of the above categories with potential MOD; for Cochrane reviews $n = 29$; for non-Cochrane reviews $n = 13$.

reviews with nonsignificant effect estimates, we restricted our survey to reviews with a statistically significant pooled effect estimates for a patient-important efficacy outcome. Our justification is that a follow-up study will use this same sample to explore the impact of different imputation methods on significant estimates [8] and that these reviews are most likely to influence clinical practice. We acknowledge that meta-analyses which properly account for missing data produce wider confidence intervals and are more likely to provide nonsignificant pooled effect estimates. Future studies should include both reviews with significant and nonsignificant results and explore the statistical significance as a potential covariate.

Finally, in general, one would expect Cochrane reviews to outperform non-Cochrane reviews regarding reporting and handling of MOD. Cochrane reviews are supposed to follow published protocols and explicitly declare their approach for dealing with MOD in meta-analysis and for assessing risk of bias associated with MOD for each trial; these standard reporting requirements as part of Cochrane reviews are not prerequisites for non-Cochrane reviews.

4.3. Comparison with similar studies

We previously examined how 202 Cochrane and non-Cochrane systematic reviews of trials published in 2010

Table 3. Handling of missing outcome data in the primary analyses of 100 Cochrane and non-Cochrane systematic reviews

	Overall ($N = 100$)	Cochrane SR ($N = 50$)	Non-Cochrane SR ($N = 50$)	*P* value [a]
Explicitly stated specific analytical method(s) for handling with MOD				
Using complete case analysis	5 (5%)	2 (4%)	3 (6%)	0.95
Assuming no participants with MOD had the event	4 (4%)	3 (6%)	1 (2%)	
Assuming all participants with MOD had the event	2 (2%)	1 (2%)	1 (2%)	
Assuming participants with MOD had same event rate as those followed up in respective randomization groups	1 (1%)	1 (2%)	0 (0%)	
Using worst case scenario[b]	1 (1%)	1 (2%)	0 (0%)	
Using best case scenario[c]	0 (0%)	0 (0%)	0 (0%)	
Using other assumption(s)	2 (2%)	1 (2%)	1 (2%)	
Using whatever assumptions the included trials used	0 (0%)	0 (0%)	0 (0%)	
Excluding trials with high rate of MOD	1 (1%)	0 (0%)	1 (2%)	
Other	3 (3%)	1 (2%)	2 (4%)	
No method described	81 (81%)	40 (80%)	41 (82%)	
Provided justification for the analytical method(s) used to handle MOD				
Yes	2 (2%)	1 (2%)	1 (2%)	0.878
No, MOD was handled but not justified	12 (12%)	7 (14%)	5 (10%)	
Not applicable, MOD was not handled in the first place	86 (86%)	42 (84%)	44 (88%)	

Abbreviation: MOD, missing outcome data; SR, systematic review.
[a] *P*-value for the difference between Cochrane and non-Cochrane systematic reviews.
[b] Worst case scenario: assuming all participants with MOD in the intervention group had the event but none in thecontrol group did.
[c] Best case scenario: assuming that all participants with MOD in the control group had the event but none in the intervention group did.

reported, handled, and assessed the risk of bias associated with MOD [6]. We found that 25% of systematic reviews reported plans for handling different categories of MOD, consistent with the 19% corresponding value in this present study. Our earlier study also found that only 6% of reviews planned sensitivity analyses to test the robustness of the results for categorical data, consistent with the 9% in this present survey. Relative to our earlier work, the present survey includes a more recent sample of reviews, assesses the explicit plans to consider various categories of participants as having MOD reported in the methods section, and inquired about taking uncertainty into account. As found in our earlier study, compared with non-Cochrane reviews, Cochrane reviews were somewhat more rigorous in their consideration of MOD [15,16].

Another systematic survey by Spineli et al [7] examined the reporting of methodology to address MOD in 190 Cochrane systematic reviews related to mental health published between 2009 and 2012. The investigators found that 16% of the eligible reviews undertook sensitivity analysis to explore the impact of MOD. The investigators also found that 79% of Cochrane systematic reviews with studies that reported MOD incorporated MOD in the primary analysis using an imputation strategy. We found that only 10% of reviews imputed outcomes of participants with MOD. A possible explanation is that mental health research suffers from MOD to a greater extent than other clinical areas of research and so investigators are much more aware of the issue [17]. It is notable that the same study found that 35% of the eligible reviews excluded from their meta-analyses trials with MOD rates greater than 50%.

4.4. Implications for conducting systematic reviews

Systematic review authors should adhere to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement recommendations relating to reporting on, and handling of, MOD in systematic reviews [18]:

- Avoid, whenever possible, imputing data when it is missing from a study report. When necessary, contact the original investigators to try to obtain missing information or confirm the data extracted with the trial authors.
- Report attempts to acquire missing information from investigators or sponsors (describe briefly who was contacted and what unpublished information was obtained);
- Report any assumptions made about MOD or unclear information and to explain those processes;
- Present study-level characteristics to clearly indicate whether any missing or unclear information exits;
- If information is imputed, state the approach that was used and for which outcomes.

Table 4. Assessing the risk of bias associated with missing outcome data in the selected meta-analyses of 100 Cochrane and non-Cochrane systematic reviews

	Overall ($N = 100$)	Cochrane SR ($N = 50$)	Non-Cochrane SR ($N = 50$)	*P* value [a]
Evaluated the risk of bias associated with MOD at the level of the trial				
Yes, using the Cochrane RoB tool	65 (65%)	43 (86%)	22 (44%)	<0.001[a]
Yes, using a tool other than the Cochrane RoB tool (e.g. Jadad's scale)	22 (22%)	5 (10%)	17 (34%)	
Not done	13 (13%)	2 (4%)	11 (22%)	
Stated method(s) used to judge risk of bias associated with MOD at the level of the meta-analysis				
Sensitivity analysis	9 (9%)	7 (14%)	2 (4%)	0.160
Subgroup analysis	0 (0%)	0 (0%)	0 (0%)	-
Other	0 (0%)	0 (0%)	0 (0%)	-
No method reported	94 (94%)	46 (92%)	48 (96%)	0.678
Provided the results of the sensitivity analysis applied to account for MOD[b]				
Yes	5 (5%)	4 (8%)	1 (2%)	0.239
No, not reported	4 (4%)	3 (6%)	1 (2%)	
Not applicable, no sensitivity analysis applied	91 (91%)	43 (86%)	48 (96%)	
Took into account the uncertainty associated with imputing outcomes in the primary or secondary analysis				
Imputed outcomes and took uncertainty into account	0 (0%)	0 (0%)	0 (0%)	0.182
Imputed outcomes but did not take uncertainty into account	10 (10%)	7 (14%)	3 (6%)	
Not applicable, no MOD or did not impute outcome	90 (90%)	43 (86%)	47 (94%)	

Abbreviations: MOD, missing outcome data; RoB, Risk of Bias; SR, systematic review.
[a] *p*-value for the difference between Cochrane and non-Cochrane systematic reviews.
[b] These results pertain to the first sensitivity analysis. Three Cochrane reviews applied a second sensitivity analysis with only one reporting its results.

Similarly, according to the Cochrane RoB tool, systematic review authors should "state whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition/exclusions where reported, and any reinclusions in analyses performed by the review authors" [19].

However, the problem is not only one of reporting but also of handling MOD. Following the Cochrane handbook recommendations [19], systematic review authors should define a priori (preferably in the protocol) a clear plan to handle MOD in the meta-analysis [19]. The Cochrane handbook refers systematic review authors to substantial literature on statistical methods for making assumptions that consider uncertainty associated with different types of imputation [2,9,10,19]. These sophisticated statistical approaches are now available for both binary [3] and continuous variables [20].

Review authors should also consider recommendations of the GRADE working group for assessing the risk of bias associated with MOD in a body of evidence [4]. For the primary meta-analysis, the GRADE working group recommends either conducting a complete case analysis or making assumptions about the outcomes of those with MOD if investigators have strong hypotheses for those outcomes. If investigators opt to make assumptions about missing observations in their primary analysis, they should consider the uncertainty associated with imputation using the appropriate statistical approaches for both binary [14] and continuous variables [15]. These approaches were developed relatively recently and require specialized software [21,22].

Systematic review authors should also be aware that intention-to-treat analysis is not a method to handle MOD but to deal effectively with noncompliance in those with available outcome data [23]. The Cochrane handbook defines the principles of intention-to-treat analyses as follows: (1) analyze participants in the intervention groups to which they were randomized, regardless of the intervention they actually received; (2) measure outcome data on all participants; and (3) include all randomized participants in the analysis [19]. Unfortunately, trialists may neither adhere to the aforementioned recommendations nor provide systematic reviewers with the needed data to confirm adherence (e.g., the number of events among those who were noncompliant and that the trialist analyzed not in their randomized arm). Review authors should clearly describe such limitations.

4.5. Implications for research

There is a need to explore approaches to judge the risk of bias associated with MOD at the level of the meta-analysis. One approach would be to evaluate the impact of different methods of handling MOD on the statistical

significance of pooled effect estimates and the associated quality of evidence. Satisfactory approaches are likely to require training and tools to facilitate their use.

Acknowledgments

We would like to acknowledge Neera Bhatnagar for her help with developing the search strategy; Dr. Reem Waziry for her help with pilot testing the screening forms; and Nada Al-Matari with acquiring full texts.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jclinepi.2018.02.016.

References

[1] Akl EA, Briel M, You JJ, Sun X, Johnston BC, Busse JW, et al. Potential impact on estimated treatment effects of information lost to follow-up in randomised controlled trials (LOST-IT): systematic review. BMJ 2012;344:e2809.

[2] Turner NL, Dias S, Ades AE, Welton NJ. A Bayesian framework to account for uncertainty due to missing binary outcome data in pairwise meta-analysis. Stat Med 2015;34:2062−80.

[3] Higgins JP, White IR, Wood AM. Imputation methods for missing outcome data in meta-analysis of clinical trials. Clin Trials 2008;5:225−39.

[4] Guyatt GH, Ebrahim S, Alonso-Coello P, Johnston BC, Mathioudakis AG, Briel M, et al. GRADE guidelines 17: assessing the risk of bias associated with missing participant outcome data in a body of evidence. J Clin Epidemiol 2017;87:14−22.

[5] Akl EA, Kahale LA, Agoritsas T, Brignardello-Petersen R, Busse JW, Carrasco-Labra A, et al. Handling trial participants with missing outcome data when conducting a meta-analysis: a systematic survey of proposed approaches. Syst Rev 2015;4:98.

[6] Akl EA, Carrasco-Labra A, Brignardello-Petersen R, Neumann I, Johnston BC, Sun X, et al. Reporting, handling and assessing the risk of bias associated with missing participant data in systematic reviews: a methodological survey. BMJ Open 2015;5(9):e009368.

[7] Spineli LM, Pandis N, Salanti G. Reporting and handling missing outcome data in mental health: a systematic review of Cochrane systematic reviews and meta-analyses. Res Synth Methods 2015;6:175−87.

[8] Akl EA, Kahale LA, Agarwal A, Al-Matari N, Ebrahim S, Alexander PE, et al. Impact of missing participant data for dichotomous outcomes on pooled effect estimates in systematic reviews: a protocol for a methodological study. Syst Rev 2014;3:137.

[9] White IR, Higgins JP, Wood AM. Allowing for uncertainty due to missing data in meta-analysis−part 1: two-stage methods. Stat Med 2008;27:711−27.

[10] White IR, et al. Allowing for uncertainty due to missing data in meta-analysis—Part 2: hierarchical models. Stat Med 2008;27:728−45.

[11] Gamble C, Hollis S. Uncertainty method improved on best-worst case analysis in a binary meta-analysis. J Clin Epidemiol 2005;58:579−88.

[12] Mavridis D, Chaimani A, Efthimiou O, Leucht S, Salanti G. Addressing missing outcome data in meta-analysis. Evid Based Ment Health 2014;17(3):85−9.

[13] Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159−74.

[14] Ebrahim S, Akl EA, Mustafa RA, Sun X, Walter SD, Heels-Ansdell D, et al. Addressing continuous data for participants excluded from trial analysis: a guide for systematic reviewers. J Clin Epidemiol 2013;66:1014−1021.e1.

[15] Hopewell S, Boutron I, Altman DG, Ravaud P. Incorporation of assessments of risk of bias of primary studies in systematic reviews of randomised trials: a cross-sectional study. BMJ Open 2013;3(8):e003342.

[16] Moseley AM, Elkins MR, Herbert RD, Maher CG, Sherrington C. Cochrane reviews used more rigorous methods than non-Cochrane reviews: survey of systematic reviews in physiotherapy. J Clin Epidemiol 2009;62:1021−30.

[17] Wahlbeck K, Tuunainen A, Ahokas A, Leucht S. Dropout rates in randomised antipsychotic drug trials. Psychopharmacology (Berl) 2001;155(3):230−3.

[18] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLos Med 2009;6(7):e1000100.

[19] Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. [updated March 2011]. The Cochrane Collaboration; 2011. Available at www.handbook.cochrane.org.

[20] Mavridis D, White IR, Higgins JP, Cipriani A, Salanti G. Allowing for uncertainty due to missing continuous outcome data in pairwise and network meta-analysis. Stat Med 2015;34:721−41.

[21] R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2010. Available at https://www.r-project.org/.

[22] StataCorp. Stata statistical software: release 15. College Station, TX: StataCorp LLC; 2017.

[23] Alshurafa M, Briel M, Akl EA, Haines T, Moayyedi P, Gentles SJ, et al. Inconsistent definitions for intention-to-treat in relation to missing outcome data: systematic review of the methods literature. PLoS One 2012;7:e49163.