Addressing MISSING DATA of Dichotomous Outcomes in Systematic Reviews

Lara A Kahaleh

Addressing Missing Data of Dichotomous Outcomes in Systematic Reviews

Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht
PhD thesis, Utrecht University, The Netherlands

ISBN: 978-9953-586-54-0
Author: Lara A Kahaleh
Cover design: Rita El Hachem
Lay-out: Rita El Hachem
Printed by: AUB bookstore

Addressing Missing Data of Dichotomous Outcomes in Systematic Reviews

Omgaan met ontbrekende gegevens van dichotome uitkomsten in systematische reviews

(met een samenvatting in het Nederlands)

PROEFSCHRIFT

Ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. H.R.B.M. Kummeling, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op

woensdag 27 november 2019
des middags te 4.15 uur

door

Lara Andrea Kahaleh

Geboren op 30 november 1989
te Beiroet, Libanon

Promotoren:

Prof. dr. R.J.P.M. Scholten

Prof. dr. E. Akl

Copromotor:

Dr. L. Hooft

Manuscripts based on the studies presented in this thesis

Chapter 2

Kahale LA, Diab B, Brignardello-Petersen R, Agarwal A, Mustafa RA, Kwong J, Neumann I, Li L, Lopes LC, Briel M, Busse JW, Iorio A, Vandvik PO, Alexander PE, Guyatt G, and Akl EA. Systematic reviews do not adequately report or address missing outcome data in their analyses: a methodological survey. **Journal of Clinical Epidemiology**. 2018 Jul 1;99:14-23.

Chapter 3

Kahale LA, Diab B, Khamis AM, Chang Y, Lopes LC, Agarwal A, Li L, Mustafa R, Koujanian S, Waziry R, Busse JW, Dakik A, Guyatt G, and Akl EA. Potentially missing data was considerably more frequent than definitely missing data in randomized controlled trials: A methodological survey. **Journal of Clinical Epidemiology.** 2018 Oct 6.

Chapter 4

Kahale LA, Guyatt GH, Agoritsas T, Briel M, Busse JW, Carrasco-Labra A, Khamis AM, Zhang Y, Hooft L, Scholten RJ, Akl EA. A guidance was developed to identify participants with missing outcome data in randomized controlled trials. **Journal of Clinical Epidemiology** (2019), doi: https://doi.org/10.1016/j.jclinepi.2019.07.003.

Chapter 5

Kahale LA, Diab B, Khamis AM, Chang Y, Lopes LC, Agarwal A, Li L, Mustafa R, Koujanian S, Waziry R, Busse JW, Dakik A, Hooft L, Guyatt G, Scholten RJMP, Akl EA. Potential impact of missing outcome data on treatment effects in systematic reviews: an imputation study. **BMJ** (submitted).

Chapter 6

Kahale LA, Diab B, Khamis AM, Chang Y, Lopes LC, Agarwal A, Li L, Mustafa R, Koujanian S, Waziry R, Busse JW, Dakik A, Hooft L, Guyatt G, Scholten RJMP, Akl EA. Methods of handling missing data were inconsistent across trials included in the same meta-analysis: a methodological survey. **International Journal of Epidemiology** (submitted).

Table of Contents

Chapter 1

General Introduction

Introduction

Randomized controlled trials (RCTs) are expected to produce unbiased estimates of treatment effects compared to other study designs (1). Problems in the design and conduct of RCTs may present threats to the validity of their results. In practice, achieving this goal depends on the extent to which potential sources of bias have been avoided or minimized. For several reasons, RCTs might suffer from missing data of many variables for a number of participants (2). Missing data could apply to baseline characteristics of trial participants or outcomes. In this thesis, we focus on missing outcome data for trial participants which we define as outcome data from included RCTs that are not available to the authors of systematic reviews (whether from published RCT reports or through contact with trialists). Despite persistent attempts by trial investigators to prevent missing outcome data, this phenomenon cannot be entirely eliminated and happens to be very common in RCT reports (3). Eight methodological surveys of various disciplines found that the percentage of RCTs with missing outcome data ranged from 16% to 100% (4-11). Among the RCTs that do report having participants with missing outcome data, the average percentage of those participants ranged from 6% to 34% (4-9, 11) .

Missing outcome data is not only prevalent, but it represents a serious potential source of bias - attrition bias - based on several factors (12). A recent study found that applying assumptions regarding outcomes of participants with missing outcome data could change the statistical significance of results of RCTs published in top medical journals (4). The potential effect of attrition bias is that invalid conclusions about efficacy and safety of studied interventions may be reached and ultimately impact clinical practice. If the number and characteristics of participants with missing outcome data differ between the randomized groups, then participants remaining in the study may no longer be comparable for their prognosis, leading to attrition bias (13, 14). Also, the mechanism of missingness, i.e., why values are missing and the connection of those reasons with treatment outcomes, contributes largely to attrition bias (15). The mechanism of missingness is classified into three categories (15, 16):

- Missing completely at random: the reason of missingness is related neither to participants' characteristics nor to the outcome, e.g., if a participant misses some

appointments due to scheduling difficulties. This assumption means that the group of participants who provided data is a random sample of the total population.

- Missing at random: the reason of missingness is related to participants' characteristics but not the actual outcome, e.g., primary school children are randomized to different intervention groups to reduce school-related anxiety. Younger children are less likely to provide outcome data due to their age-related cognitive challenges. Thus, rates of missing outcome data among younger children across groups are expected to be comparable, and consequently the outcomes for the younger children who dropped out are expected to be similar to outcomes for the younger children who completed the study.
- Missing not at random or informatively missing: the reason of missingness is associated with the actual effect of the intervention, e.g., in mental health trials, placebo groups show larger dropout rate than patients treated with antipsychotics because of placebo's lack of efficacy. Thus, the effect estimate of the relative treatment would be biased when the analysis is based only on participants who completed the study.

For judging whether certain participants are missing at random or not, it is very crucial to distinguish between premature end of follow-up, which is specific to a participant, and missing outcome data, which is specific to an outcome (17). Premature end of follow-up is defined as 'the cessation of following up of a specific participant before the planned end of study follow-up' whereas missing outcome data refers to 'the unavailability of data for a specific outcome for a specific participant' (17). Thus, premature end of follow-up could result in missing data for some outcomes (i.e., outcomes that occurred after the participant being lost to follow-up) but not all outcomes.

Extensive literature have been provided on (1) strategies to avoid missing outcome data at the design and study conduct level (15, 18-20), and (2) statistical methods to handle missing outcome data in RCTs (5, 6, 15, 21-29). Some methods include applying naïve approaches like assuming all or none of the participants with missing outcome data developed the outcome of interest. Another more plausible methods suggested assume that the incidence of developing the outcome among participants with missing outcome

data is relative to the incidence of developing the outcome among participants with complete follow-up.

In order to preserve the prognostic balance created by randomization, the intention-to-treat principle calls for trialists to include all randomized participants in the group to which they were allocated to (30). The CONSORT (Consolidated Standards of Reporting Trials) statement, which provides guidance to improve the quality of reporting of clinical trials, recommends intention-to-treat analysis as standard practice (31, 32). Though this principle is frequently applied, the intention-to-treat principle does not protect against bias associated with missing outcome data (30). Indeed, missing outcome data is still present in one quarter of RCT reports, and is more poorly reported than other items listed in CONSORT (33). Moreover, one would still have to make assumptions about the outcomes of participants with missing outcome data in order to include them in the analysis (34). A common practice by most trials investigators is the inclusion participants with missing outcome data in the denominators while calculating estimates of effect. This approach assumes that none of those participants with missing outcome data experienced the outcome of interest. Consequently, reporting results of the effect of the intervention may be misleading given that this assumption is highly unlikely.

The problem of missing outcome data is carried over to meta-analysis of RCTs. A recent systematic survey of 387 systematic reviews of RCTs comparing at least three interventions found that 70% of systematic reviews explicitly reported that there are missing outcome data in the included trials (35). Systematic review authors face several challenges when abstracting data related to missing outcome data from RCTs, mainly due to poor reporting of the RCTs. First, while the systematic reviewer needs missing outcome data information specific to the outcome being meta-analyzed, often RCTs report instead the number of participants with premature end of follow-up (which does not necessarily imply missing outcome data for all those participants for that outcome) (17). Second, it is not always clear whether the RCT authors followed-up participants in certain categories (e.g., withdrew consent, were non-compliant) for the outcome of interest (i.e., whether they have missing outcome data or not) (17). Third, it is not always clear whether or how the RCT authors dealt with missing outcome data in their analysis (e.g., complete case analysis versus imputing outcomes) (17). Indeed, a recent

methodological survey found that the majority of systematic reviews does not provide any strategy to address missing outcome data in their analysis (35). Fourth, very often, results of RCTs are usually presented together for fully observed and imputed outcomes which makes it hard to abstract the required data from RCT reports (16). This poor reporting of missing outcome data information in RCTs contributes to the inadequate reporting and handling of missing outcome data in systematic reviews (5, 6, 9, 13, 36-45).

A crucial issue for all authors of systematic reviews is the risk of attrition bias in included RCTs. The Cochrane Collaboration's Risk of Bias (RoB) tool was designed to help in assessing bias associated with a number of factors including incomplete outcome data (12). However, a study assessing stakeholders' experiences with and perceptions of the Cochrane RoB tool participants found that incomplete outcome data as one of the most difficult domains to assess (46). Stakeholders also requested more guidance on how to incorporate RoB assessments into meta-analyses and conclusion (46).

Aims and objectives of the work presented in the thesis

The overall aim of this thesis is to provide systematic review authors with specific guidance on how to address missing outcome data of trial participants in their reviews. The specific aims are:

1. To describe how systematic review authors report on the categories of participants who might have missing outcome data, handle missing outcome data in their primary meta-analyses of dichotomous outcomes, and assess the associated risk of bias;
2. To assess how trial authors report on the categories of participants that might have missing outcome data and their follow-up status, and on the handling of these participants in their main and secondary analyses;
3. To provide guidance for authors of systematic reviews on how to identify and classify participants with missing outcome data in trials;
4. To assess risk of bias associated with missing outcome data in systematic reviews by examining how different methods of handling missing outcome data alter statistical significance of pooled effect estimates of dichotomous outcomes and

quantifying the change in effect estimate when applying different methods of handling missing outcome data;

5. To assess whether systematic review authors are consistent in their methods of handling missing outcome data across trials included in their meta-analyses, and whether the methods used for handling missing outcome data in their meta-analyses were consistent with the reported methods.

Thesis outline

Chapter 2 is a methodological survey of 100 systematic reviews to explore how systematic review authors report and address categories of trial participants with potential missing data of dichotomous outcomes. The study focuses on dichotomous outcome data, given the methodological and statistical issues vary substantively for continuous data. Chapter 3 is a methodological survey of all RCTs that were included in the 100 systematic reviews of Chapter 2. We assess how trial authors (1) report on different categories of participants that might have missing outcome data, (2) handle these categories in the analysis, and (3) judge the risk of bias associated with missing outcome data. In chapter 4, we develop guidance for systematic review authors on identifying participants with missing outcome data in RCTs. Guidance statements are informed by a review of studies addressing the topic of missing outcome data and an iterative process of feedback and refinement, through meetings involving experts in health research methodology and authors of systematic reviews. Chapter 5 is an imputation study, which assesses the risk of bias associated with missing outcome data in systematic reviews. Specifically, we examine how different methods of handling missing outcome data alter statistical significance of pooled effect estimates of dichotomous outcomes. Also, we quantify the change in effect estimate when applying different methods of handling missing outcome data. Last, in chapter 6, we explore how authors of the 100 systematic reviews (same sample as chapter 2) actually dealt with trial missing outcome data in the meta-analysis. We assess whether these methods are consistently applied across all trials included in the meta-analysis. When consistent, we check whether these methods are consistent with the methods reported in chapter 2. In chapter 7, we describe the strengths and limitations of the whole project, implications for

practice for (1) systematic review authors, (2) trialists, (3) journal editors, and implications for research.

References

1. Wood, A.M., I.R. White, and S.G. Thompson, *Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals.* Clin Trials, 2004. **1**(4): p. 368-76.
2. Akl, E.A., 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**: p. e2809.
3. Altman, D.G., *Missing outcomes in randomized trials: addressing the dilemma.* Open Med, 2009. **3**(2): p. e51-3.
4. Akl, E.A., 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**: p. e2809.
5. Bell, M.L., et al., *Handling missing data in RCTs; a review of the top medical journals.* BMC Med Res Methodol, 2014. **14**: p. 118.
6. Fielding, S., et al., *Reporting and dealing with missing quality of life data in RCTs: has the picture changed in the last decade?* Qual Life Res, 2016. **25**(12): p. 2977-2983.
7. Hussain, J.A., et al., *Quality of missing data reporting and handling in palliative care trials demonstrates that further development of the CONSORT statement is required: a systematic review.* J Clin Epidemiol, 2017. **88**: p. 81-91.
8. Powney, M., et al., *A review of the handling of missing longitudinal outcome data in clinical trials.* Trials, 2014. **15**(1): p. 237.
9. Ibrahim, F., et al., *A systematic review of randomised controlled trials in rheumatoid arthritis: the reporting and handling of missing data in composite outcomes.* Trials, 2016. **17**(1): p. 272.
10. Shivasabesan, G., B. Mitra, and G.M. O'Reilly, *Missing data in trauma registries: A systematic review.* Injury, 2018. **49**(9): p. 1641-1647.
11. Sullivan, T.R., et al., *Treatment of missing data in follow-up studies of randomised controlled trials: A systematic review of the literature.* Clin Trials, 2017. **14**(4): p. 387-395.
12. The Cochrane Collaboration, *Cochrane Handbook for Systematic Reviews of Interventions*. Vol. Version 5.1.0. 2011.
13. Wood, A.M., I.R. White, and S.G. Thompson, *Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals.* Clinical trials, 2004. **1**(4): p. 368-376.
14. Dumville, J.C., D.J. Torgerson, and C.E. Hewitt, *Reporting attrition in randomised controlled trials.* BMJ, 2006. **332**(7547): p. 969-71.
15. National Research Council, *The Prevention and Treatment of Missing Data in Clinical Trials*. 2010, The National Academic Press: Washington (DC).
16. Mavridis, D., et al., *Addressing missing outcome data in meta-analysis.* Evid Based Ment Health, 2014. **17**(3): p. 85-9.
17. Akl, E.A., et al., *Three challenges described for identifying participants with missing data in trials reports, and potential solutions suggested to systematic reviewers.* J Clin Epidemiol, 2016.
18. Brueton, V.C., et al., *Strategies to improve retention in randomised trials.* The Cochrane database of systematic reviews, 2013(12): p. 1.
19. Dziura, J.D., et al., *Strategies for dealing with missing data in clinical trials: from design to analysis.* Yale J Biol Med, 2013. **86**(3): p. 343-58.

20. Hollestein, L. and J. Carpenter, *Missing data in clinical research: an integrated approach.* British Journal of Dermatology, 2017. **177**(6): p. 1463-1465.
21. Bell, M.L. and D.L. Fairclough, *Practical and statistical issues in missing data for longitudinal patient-reported outcomes.* Stat Methods Med Res, 2014. **23**(5): p. 440-59.
22. Carpenter, J. and M. Kenward, *Guidelines for handling missing data in social science research.* 2007.
23. Fielding, S., et al., *Methods for handling missing data in palliative care research.* Palliative medicine, 2006. **20**(8): p. 791-798.
24. Fielding, S., P. Fayers, and C.R. Ramsay, *Analysing randomised controlled trials with missing data: choice of approach affects conclusions.* Contemp Clin Trials, 2012. **33**(3): p. 461-9.
25. Jakobsen, J.C., et al., *When and how should multiple imputation be used for handling missing data in randomised clinical trials–a practical guide with flowcharts.* BMC medical research methodology, 2017. **17**(1): p. 162.
26. Jorgensen, A.W., et al., *Comparison of results from different imputation techniques for missing data from an anti-obesity drug trial.* PLoS One, 2014. **9**(11): p. e111964.
27. Li, P., E.A. Stuart, and D.B. Allison, *Multiple imputation: a flexible tool for handling missing data.* Jama, 2015. **314**(18): p. 1966-1967.
28. O'Neill, R.T. and R. Temple, *The prevention and treatment of missing data in clinical trials: an FDA perspective on the importance of dealing with it.* Clin Pharmacol Ther, 2012. **91**(3): p. 550-4.
29. Preston, N.J., et al., *Recommendations for managing missing data, attrition and response shift in palliative and end-of-life care research: part of the MORECare research method guidance on statistical issues.* Palliative medicine, 2013. **27**(10): p. 899-907.
30. Montori, V. and G. Guyatt, *Intention-to-treat principle.* Canadian Medical Association Journal, 2001. **165**(10).
31. Rennie, D., *CONSORT revised--improving the reporting of randomized trials.* JAMA, 2001. **285**(15): p. 2006-7.
32. Moher, D., et al., *CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials.* Int J Surg, 2012. **10**(1): p. 28-55.
33. Sylvestre, Y., *CONSORT: missing missing data guidelines, the effects on HTA monograph reporting.* Trials, 2011. **12**(Suppl 1): p. A61.
34. Alshurafa, M., et al., *Inconsistent definitions for intention-to-treat in relation to missing outcome data: systematic review of the methods literature.* PLoS One, 2012. **7**(11): p. e49163.
35. Spineli, L.M., J. Yepes-Nuñez, and H. Schünemann, *A systematic survey shows that reporting and handling of missing outcome data in networks of interventions is poor.* 2018.
36. Hussain, J.A., et al., *Missing data in randomised controlled trials evaluating palliative interventions: a systematic review and meta-analysis.* The Lancet, 2016. **387**: p. S53.
37. Adewuyi, T.E., G. MacLennan, and J.A. Cook, *Non-compliance with randomised allocation and missing outcome data in randomised controlled trials evaluating surgical interventions: a systematic review.* BMC Res Notes, 2015. **8**: p. 403.
38. Akl, E.A., et al., *Reporting missing participant data in randomised trials: systematic survey of the methodological literature and a proposed guide.* BMJ Open, 2015. **5**(12): p. e008431.
39. Fielding, S., et al., *A review of RCTs in four medical journals to assess the use of imputation to overcome missing data in quality of life outcomes.* Trials, 2008. **9**: p. 51.

40. Fiero, M.H., et al., *Statistical analysis and handling of missing data in cluster randomized trials: a systematic review.* Trials, 2016. **17**: p. 72.
41. Gewandter, J.S., et al., *Reporting of missing data and methods used to accommodate them in recent analgesic clinical trials: ACTTION systematic review and recommendations.* Pain, 2014. **155**(9): p. 1871-7.
42. Karlson, C.W. and M.A. Rapoff, *Attrition in randomized controlled trials for pediatric chronic conditions.* J Pediatr Psychol, 2009. **34**(7): p. 782-93.
43. Masconi, K.L., et al., *Reporting and handling of missing data in predictive research for prevalent undiagnosed type 2 diabetes mellitus: a systematic review.* EPMA J, 2015. **6**(1): p. 7.
44. Spineli, L.M., *Missing binary data extraction challenges from Cochrane reviews in mental health and Campbell reviews with implications for empirical research.* Res Synth Methods, 2017. **8**(4): p. 514-525.
45. Wahlbeck, K., et al., *Dropout rates in randomised antipsychotic drug trials.* Psychopharmacology, 2001. **155**(3): p. 230-233.
46. Savovic, J., et al., *Evaluation of the Cochrane Collaboration's tool for assessing the risk of bias in randomized trials: focus groups, online survey, proposed recommendations and their implementation.* Syst Rev, 2014. **3**: p. 37.

Chapter 2

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

Kahale LA, Diab B, Brignardello-Petersen R, Agarwal A, Mustafa RA, Kwong J, Neumann I, Li L, Lopes LC, Briel M, Busse JW, Iorio A, Vandvik PO, Alexander PE, Guyatt G, and Akl EA
Journal of clinical epidemiology. 2018 Jul 1; 99:14-23

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=534 in methods section and n= 56 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=59) 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.

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]. The GRADE working group 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-up. 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.

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.

Methods

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.

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.

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 https://www.nlm.nih.gov/bsd/aim.html). The search included studies published in 2012, in any language. Appendix on the journal's website at www. Elsevier.com provides the details of our search strategy.

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.

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, we 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 analyses. 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].

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 less than five. 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).

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).

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.

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'' (21%) and ''explained LTFU'' (6%). 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.

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

Assessing risk of bias associated with MOD

Table 4 describes the assessment of the risk of bias associated with MOD. 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-value <0.001). Of the nine reviews that reported 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.

Discussion

Summary of findings

Although 42 of 100 systematic reviews reported on at least one of the 10 pre-defined categories of MOD, only 19 reported plans for handling MOD in their analyses. The majority of reviews (87%) addressed risk of bias associated with MOD at the trial level, however, a small percentage (9%) reported 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.

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, and conducting calibration exercises. Furthermore, we explored issues that have not previously been addressed, particularly in terms of categorizing MOD. Up to 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 reviews with non-significant 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 non-significant pooled effect estimates. Future studies should include both reviews with significant and non-significant 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.

Comparison with similar studies

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

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 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.

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 re-inclusions 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 by 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 non-compliance 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 non-compliant and that the trialist analyzed not in their randomized arm). Review authors should clearly describe such limitations.

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.

Figures, tables, and supplementary data

Legends

Appendix: Search strategy for non-Cochrane reviews, using Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present>

Figure 1: PRISMA study flow diagram.

Appendix 1: List of journals

Table 1: General characteristics of included systematic reviews

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

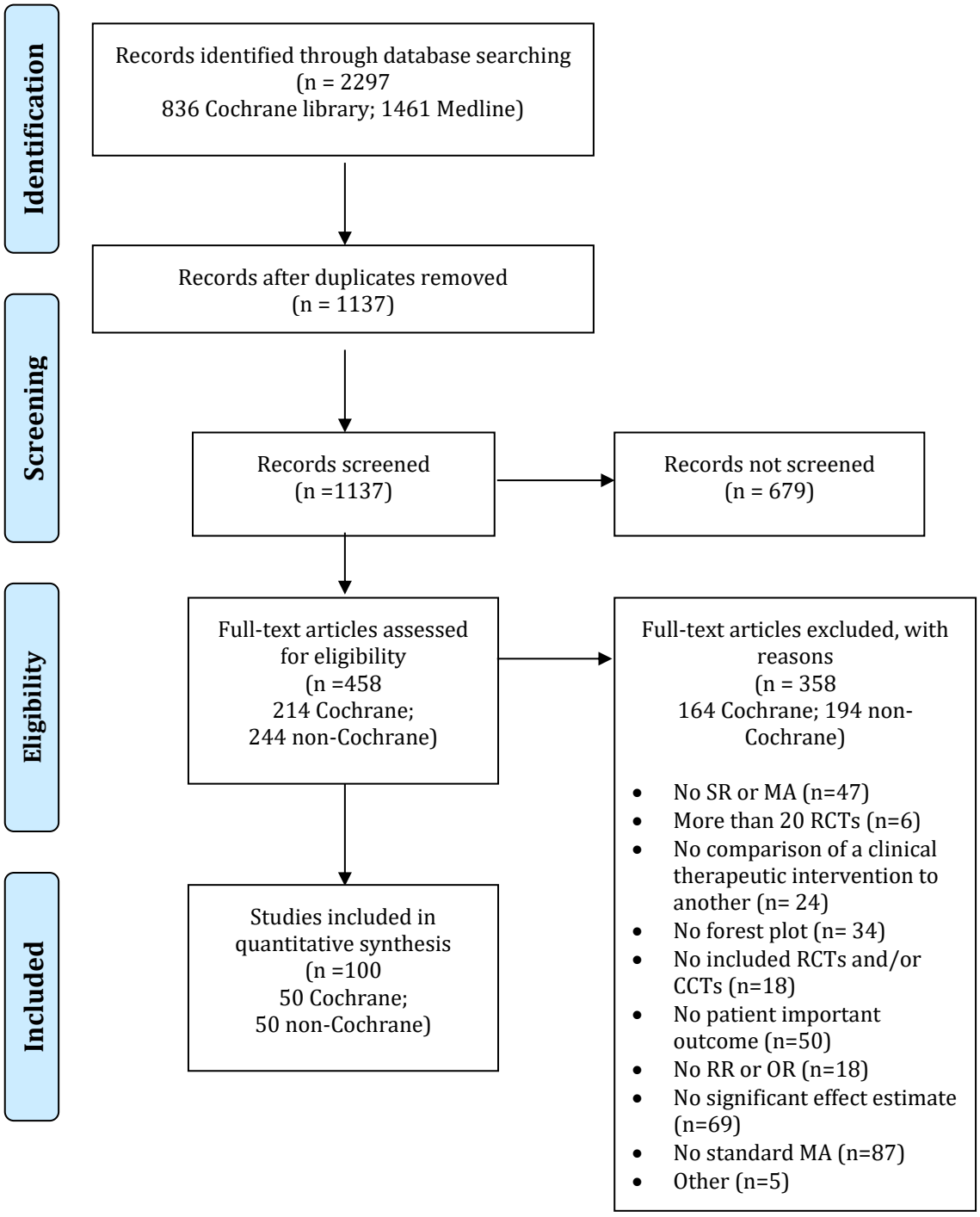
Table 3: Handling of missing outcome data in the primary analyses of 100 Cochrane and non-Cochrane systematic reviews

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

Appendix: Search strategy for non-Cochrane reviews, using Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present>

1 meta analysis.pt.
2 meta anal$.mp.
3 metaanal$.mp.
4 metanal$.mp.
5 meta epidemiolog*.mp.
6 systematic review$.mp.
7 systematic overview$.mp.
8 ((pool: or combined or combining) adj (data or trial* or studies or results)).mp.
9 ((hand adj2 search:) or handsearch:).mp.
10 cochrane.mp.
11 ((quantitative or systematic: or methodologic: or integrative:) adj2 (review: or overview: or synthes: or survey:)).mp.
12 (peto or der simonian or dersimonian).mp.
13 or/1-12
14 (pooled analys: or pooling or mantel haenszel:).mp.
15 fixed effect:.mp.
16 (extraction or medline or embase or pubmed or cinahl).ab.
17 14 or 15 or 16
18 (review: or cochrane).mp.
19 16 and 17
20 13 or 19
21 aim.sb.
22 '2012'.yr.
23 20 and 21 and 22

Figure 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



Appendix 1: List of journals

Journal Name	Journal ID	Number of reviews included
American heart journal	4	1
The American journal of cardiology	5	6
The American journal of medicine	8	2
American journal of obstetrics and gynecology	10	1
American journal of surgery	17	1
Anesthesia and analgesia	21	1
Annals of internal medicine	24	4
Annals of surgery	26	5
BJOG : an international journal of obstetrics and gynecology	34	1
BMJ (Clinical research ed.)	36	4
The British journal of surgery	40	2
Chest	43	2
Circulation	44	1
Critical care medicine	50	2
Digestive diseases and sciences	53	1
Heart (British Cardiac Society)	59	2
JAMA internal medicine	64	2
JAMA pediatrics	68	2
The Journal of bone and joint surgery. American volume	72	4
Journal of the American College of Cardiology	86	1
The Journal of thoracic and cardiovascular surgery	88	2
The Journal of trauma and acute care surgery	89	1
Obstetrics and gynecology	103	1
Pediatrics	106	1
Cochrane Database of Systematic Reviews	120	50

Table 1: General characteristics of included systematic reviews

	Overall (N=100)	**Cochrane SR (N=50)**	**Non-Cochrane SR (N=50)**	**p-value**
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.

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
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%)	0.50
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%)	0.50
Discontinued trial prematurely	0 (0%)	0 (0%)	0 (0%)	-
Cross-over	0 (0%)	0 (0%)	0 (0%)	-
Dead	1 (1%)	1 (2%)	0 (0%)	0.50
Adverse events	1 (1%)	0 (0%)	1 (2%)	0.50
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%)	1.00
Cross-over	1 (1%)	1 (2%)	0 (0%)	0.50
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

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
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%)	

Abbreviations: **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 the control 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.

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
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.

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;34:2062e80.
3. Higgins JP, White IR, Wood AM. Imputation methods for missing outcome data in meta-analysis of clinical trials. Clin Trials 2008;5: 225e39.
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;87:14e22.
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: 175e87.
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;27:711e27.
10. White IR, et al. Allowing for uncertainty due to missing data in metaanalysis -Part 2: hierarchical models. Stat Med 2008;27:728e45.
11. Gamble C, Hollis S. Uncertainty method improved on best-worst case analysis in a binary meta-analysis. J Clin Epidemiol 2005;58: 579e88.
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):85e9.
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159e74.
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;66:1014e1021.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;62:1021e30.
17. Wahlbeck K, Tuunainen A, Ahokas A, Leucht S. Dropout rates in randomised antipsychotic drug trials. Psychopharmacology (Berl) 2001;155(3):230e3.

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;34:721e41.
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.

Chapter 3

Potentially missing data was considerably more frequent than definitely missing data in randomized controlled trials: A methodological survey

Kahale LA, Diab B, Khamis AM, Chang Y, Lopes LC, Agarwal A, Li L, Mustafa R, Koujanian S, Waziry R, Busse JW, Dakik A, Guyatt G, and Akl EA.
Journal of Clinical Epidemiology. 2018 Oct 6.

Abstract

Background and Objective:

Missing data for the outcomes of participants in randomized controlled trials (RCTs) are a key element of risk of bias assessment. However, it is not always clear from RCT reports whether some categories of participants were followed-up or not (i.e., do or do not have missing data) nor how the RCT authors dealt with missing data in their analyses. Our objectives were to describe how RCT authors (1) report on different categories of participants that might have missing data, (2) handle these categories in the analysis, and (3) judge the risk of bias associated with missing data.

Methods:

We surveyed all RCT reports included in 100 clinical intervention systematic reviews (SRs), half of which were Cochrane SRs. Eligible SRs reported a group-level meta-analysis of a patient-important dichotomous efficacy outcome, with a statistically significant effect estimate. Eleven reviewers, working in pairs, independently extracted data from the primary RCT reports included in the SRs. We pre-defined 19 categories of participants that might have missing data. Then, we classified these participants as follows: 'explicitly followed- up', 'explicitly not followed-up' (i.e., definitely missing data), or 'unclear follow-up status' (i.e., potentially missing data).

Results:

Of 638 eligible RCTs, 400 (63%) reported on at least one of the pre-defined categories of participants that might have missing data. The median percentage of participants who were explicitly not followed-up was 5.8% (interquartile range 2.2-14.8%); it was 9.7% (4.1-14.9%) for participants with unclear follow-up status; and 11.7% (interquartile range 5.6-23.7%) for participants who were explicitly not followed-up and with unclear follow-up status. When authors explicitly reported not following-up participants, they most often conducted complete case analysis (54%). Most RCTs neither reported on missing data separately for different outcomes (99%) nor reported using a method for judging risk of bias associated with missing data (95%).

Conclusion:

'Potentially missing data' are considerably more frequent than 'definitely missing data'. Adequate reporting of missing data will require development of explicit standards on which editors insist and to which RCT authors adhere.

Introduction

Missing data for the outcomes of participants may threaten the validity of results of randomized controlled trials (RCTs) and systematic reviews (SRs) that include those RCTs [1]. Although missing data reduce power because of participant losses and likely increases the risk of bias particularly when missingness is associated with the occurrence of the outcome, limitations in its reporting and analysis further undermine validity. A survey of more than 200 RCTs showed that substantial discrepancy exists between proposed methodologies and current practice in handling, analysis, and reporting of missing data for patient-reported outcome measures in RCTs [2].

Trial authors typically report missing data by participant and not by outcome [3]. Consequently, when RCT authors report that a participant is lost to follow-up, SR authors might conclude that data are missing for all outcomes of interest. Nevertheless, it is possible that this participant might have experienced an event for a certain outcome before the date of premature end of follow-up. When this RCT contributes to a meta-analysis, this scenario could introduce bias whether the SR authors conduct a complete case analysis or make assumptions regarding the different outcomes of the participants with missing data.

Also, RCT authors may not report clearly whether they followed-up participants in certain categories, e.g., whether non-adherent participants did or did not have missing data [3]. Although RCT authors should follow-up non-adherent participants and use their outcome information in the analysis [4], many wrongly equate non-adherence with missing data [5].

Marciniak et al. compared the loss to follow-up rates in published reports of RCTs of oral antithrombotic agents with loss to follow-up rates calculated based on more detailed documents made available to the Food and Drug Administration (FDA) for the same RCTs [6]. They found a large discrepancy between the median of published rate of all ‘‘missing follow-up categories’’ (0.9%), and the median of the FDA-calculated loss to follow-up rates (13%). This suggests that missing data might be more frequent than what is explicitly reported and published.

It is also frequently unclear how RCT authors handle missing data in their analyses. A survey of RCTs published in three major pain journals between 2006 and 2012 found that only 45% of RCTs reported a statistical method to handle missing data in the primary meta-analysis [7]. Another review of RCTs reporting eight widely used patient-reported outcome measures found that almost half of the RCTs did not report a method to handle missing data, three-quarters did not perform sensitivity analyses, and even fewer (16%) discussed the potential impact of missing data on their results [2].

Objectives

The objective of this study was to assess the reported extent and handling of missing dichotomous outcome data in RCTs. Specifically, we examined whether and how RCT authors report on (1) the number of participants belonging to different categories that might have missing data, (2) the explicit reporting on the follow-up status of these participants, and (3) the handling of these participants in the main and secondary analyses.

Methods

Design overview

The current study is part of a larger project examining methodological issues related to missing data in SRs and RCTs [8]. We included reports of all RCTs that contributed to the main meta-analysis of the comparison and outcome addressed in a random sample of 100 eligible SRs. An eligible SR was either a Cochrane or a non-Cochrane SR published in 2012 reporting a group-level meta-analysis of a patient-important dichotomous efficacy outcome, with a statistically significant effect estimate.

Based on previous work, we developed a list of pre-defined categories of participants that might have missing data [1, 8, 9]. We refined the original list to accommodate new categories that emerged from data abstraction and did not fit existing categories. The labeling of these categories reproduces the wording used by the trial authors. Figure 1 lists the 19 final categories of participants that might have missing data (first

column) and the reported follow-up status of participants (second column). To avoid confusion among categories with common words, we defined the following categories as follows:

- 'Ineligible participants or mistakenly randomized': participants who were discovered to be ineligible after randomization, but the reason for ineligibility relates to a baseline characteristic;
- 'Ineligible because of occurrence of outcome': participants who were eligible at baseline then developed the outcome at early stages of the trial. These are typically considered ineligible if the trialists judge that the occurrence of the outcome cannot be related to the intervention of interest;
- 'Discontinued because of adverse events': participants who had adverse event and 'discontinued' either the medication or the trial;
- 'Experienced adverse events': participants who developed side effects but without any indication of 'discontinuation'.

We did not consider participants described as 'dead' and ''excluded as part of center exclusion'' as categories that might have missing data. Our published protocol includes further details regarding definitions, eligibility criteria, search strategy, and selection process of SRs [8]. Because it did not involve human participants, no ethical approval was required for the conduct of this study.

Data abstraction

We developed and pilot tested a standardized data abstraction form that included detailed instructions. We conducted calibration exercises to verify the accuracy and consistency of the data abstraction process. Eleven reviewers, working in pairs and independently, abstracted data using the REDCap electronic data capture tool hosted at the American University of Beirut [10]. A core team (E.A.A., L.A.K., B.D., and A.D.) met regularly to discuss the progress and challenges encountered during data abstraction and suggested solutions that they communicated to the entire team. They conducted triplicate and independent data abstraction as needed to ensure the quality of the data.

As presented in the following, we collected the following data from all included RCT reports: (1) characteristics of the RCTs; (2) reporting on and handling of the pre-defined

categories of missing data; (3) reporting on and handling of missing data as defined by RCT authors; and (4) assessment of risk of bias associated with missing data.

A. The characteristics of the RCTs:
   - Type of article (i.e., abstract, full-text article);
   - Language of report;
   - Type of source of funding;
   - Planned follow-up time for outcome of interest;
   - Time-to-occurrence for outcome of interest;
   - Number of participants randomized to each study arm;
   - Type of analysis (e.g., intention-to-treat [ITT], per protocol).

B. The reporting on and handling of the pre-defined categories of participants that might have missing data:
   - Reporting on participants belonging to each category in the results section, if applicable;
   - Number of participants belonging to each category;
   - Explicit reporting on follow-up status of participants within each category. We classified these participants as follows: 'explicitly followed-up', 'explicitly not followed-up' (i.e., definitely missing data), or 'unclear follow-up status' (i.e., potentially have missing data);
   - Inclusion of participants of each category in the denominator of the analysis of interest;
   - Explicit statement of the analytical method for handling each category in the analysis of interest (i.e., when generating the best effect estimate).

C. The reporting on and handling of missing data as defined by RCT authors:
   - Explicit reporting on missing data in the results section for the specific outcome of the analysis of interest as opposed to reporting premature end of follow-up for RCT participants in general [3];
   - Reporting the level of missing data (e.g., per arm, both arms combined);

- Comparison of the baseline characteristics of participants with and without missing data (e.g., missing data group vs. non-missing data group; missing data in intervention arm vs. missing data in control arm);
- Comparison of the number of participants with missing data between the two study arms;
- Description of the mechanism of missingness (e.g., missing at random);
- Explicit statement taking into account uncertainty associated with imputing outcomes when calculating the confidence interval, in case imputation methods were used in the analysis of interest;
- Justification for the analytical method used to handle missing data in the analysis of interest.

D. The assessment of the risk of bias associated with missing data:
   - Method(s) used to judge risk of bias associated with missing data (sensitivity analysis, e.g., complete case analysis, assumptions).

We also asked the data abstractors about their perception of the clarity of reporting of missing data in the included RCTs. We did not follow strict criteria to complete this answer, but the judgment was driven by the amount of time and effort spent to abstract information on all the previous variables.

Data analysis

We conducted a descriptive analysis of all variables. For categorical variables, we reported frequencies and percentages. For continuous variables, which were not normally distributed, we used median and interquartile range (IQR). We calculated the percentage of RCTs that reported on each category of participants that might have missing data. In addition, for each category, we calculated the percentage of RCTs that (1) explicitly reported following-up participants, (2) explicitly reported not following-up participants, and (3) did not provide explicit reporting on the follow-up status. We created a variable called 'all categories combined' which includes for each RCT, the participants that belonged to at least 1 of the 19 categories.

Also, we calculated the percentage of participants belonging to each of the 19 categories of participants separately and for all categories combined. We then calculated the distribution (median and IQR) of these percentages across RCTs.

To assess the potential impact of missing data on study effect estimates, we calculated for each RCT that reported on at least one category of participants that might have missing data, the ratio of the percentage of participants with missing data to the difference in the event rates (denominator being number randomized) between the two arms. We clarify the calculation in an example in Supplementary File. We then calculated the median and IQR for the distribution of these ratios across RCTs. We made these calculations twice: first, considering participants who were explicitly not followed-up; second, considering both participants who were explicitly not followed-up and those with unclear follow-up status. For all analyses, we used SPSS statistical software, V.21.0 (SPSS INC, Chicago, IL, USA).

Results

Of 653 RCTs included in the 100 eligible SRs, we could not retrieve the full-texts for 15 (2.3%), despite extensive efforts by librarians. Table 1 reports the general characteristics of the 638 included RCTs. The median date of publication was 2005. Most were published in English (94%), reported source of funding (55%), reported planned follow-up time (80%), assessed a pharmacologic intervention (56%), but did not report type of analysis (e.g., ITT, per protocol; 61%). The median planned follow-up time of the outcome of interest was 6 months, and the median time-to-occurrence was 1 month. The median numbers of participants randomized to the intervention and control arms were 69 (31-167) and 66 (32-162), respectively.

Reporting on and handling of the pre-defined categories of missing data

Table 2 shows the reporting, in the results section, on each of the pre-defined categories of missing data (please refer to Fig. 2 for explanation of what Table 2 reports).

The top reported categories were ‘‘unexplained lost to follow-up’’ (25%) and ‘‘ineligible or mistakenly randomized’’ (22%). The least reported categories were ‘‘protocol violation by investigator or clinician’’ (0.6%) and ‘‘lack of efficacy’’ (1.3%). Considering

the categories separately, only for the one category ‘‘experienced adverse events,’’ a majority of the RCTs (68%) explicitly reported following-up participants. For the remaining categories excluding those who were explicitly followed-up, most RCTs did not explicitly report on whether they followed-up participants (range 60-86%).

The distribution of all 638 RCTs according to their reporting on categories of participants that might have missing data was as follows:

- One hundred eighty-seven RCTs (29%) did not report on any category of participants that might have missing data;
- Fifty-one RCTs (8%) reported on at least one category of participants who were explicitly followed-up;
- Four hundred RCTs (63%) reported on at least one category of participants that were either explicitly not followed-up or with unclear follow-up status.

Table 3 shows the distribution of percentages of participants across the RCTs belonging to each category according to follow-up status (please refer to Fig. 2 for explanation of what Table 3 reports). The categories with the largest median percentages of participants were ‘outcome not assessable’ (27.3%, IQR 3.7-28.7%) and ‘experienced adverse events’ (18.3%, IQR 7.2-46.2%). The three categories with the smallest median percentages of participants were ‘withdrawn by investigator or clinician’ (1.3%, IQR 0.6-4.2%), ‘did not receive the first dose’ (1.4%, IQR 0.9-3.1%), and ‘unintended protocol violation’ (1.4%, IQR 0.3-4.9%).

Among 256 RCTs that explicitly reported on participants who were explicitly not followed-up for at least one of the pre-defined categories, the median percentage for all categories combined was 5.8% (IQR 2.2-14.8%).

Among 288 RCTs that mentioned at least one of the pre-defined categories of participants with unclear follow-up status, the median percentage for all categories combined was 9.7% (IQR 4.1-19.9%). Among the 400 RCTs that reported on participants that were either explicitly not followed-up or with unclear follow-up status for at least one of the pre-defined categories, the median percentage for all categories combined was 11.7% (IQR 5.6-23.7%).

Table 4 shows the handling in the primary analysis of each of the pre-defined categories of participants that were either explicitly not followed-up or with unclear follow-up status by the 400 RCTs that reported on at least one of those categories. We present these handling methods according to the follow-up status of these participants. Each of the 400 RCT could have reported on more than one category, resulting in a total of 998 instances where categories were mentioned:

- In the 362 instances in which the authors explicitly reported not following-up participants, 54% reported using complete case analysis, 43% did not report how they dealt with missing data, and 3% reported using another specific method (e.g., none had the event, all had the event, multiple imputation);
- In the 636 instances in which the authors did not explicitly report on the follow-up status of participants, 70% did not report how they dealt with participants with potential missing data, 29% reported using complete case analysis, and 1% reported using another specific method.

Reporting on and handling of missing data as defined by RCT authors

Table 5 shows the reporting on, handling of, and assessing risk of bias associated with missing data as defined by the RCT author. These participants who might have missing data include both participants who were explicitly not followed-up and participants with unclear follow-up status. Among the 400 RCTs that reported at least one category of participants that might have missing data in the results section, the majority reported the number of participants with missing data per arm (88%). However, a minority of the RCTs reported on missing data separately for different outcomes (1%), compared the baseline characteristics of participants with and without missing data separately for each study arm (2%) or of participants with missing data separately between the two study arms (1%), or compared the number of participants with missing data separately between the two study arms (6%). Only three studies (0.7%) described mechanisms of missingness (e.g., missing completely at random, missing not at random) of participants with missing data. None of the 13 RCTs that imputed outcomes took uncertainty into account when calculating the confidence interval. Only, three RCTs presented a justification for their approach to handle missing data (0.5%). In addition, 95% did not report using a method for judging risk of bias associated with missing data.

Clarity of reporting of missing data

Reviewers' judgment on whether the reporting of missing data was clear was as follows: 34% agreed, 36% were neutral, and 30% disagreed. The agreement between the pairs of reviewers on those judgments was good (Kappa = 0.731). Box 1 lists examples of the challenges that the reviewers faced during data abstraction and management and how they addressed them.

Ratio of rate of missing data to the risk difference

The ratio of the rate of missing data relative to the risk difference for participants who were explicitly not followed-up was median 0.6, IQR 0.0-3.0. When included in this analysis both participants who were explicitly not followed-up and those with unclear follow-up status, the ratio rises to a median of 1.7, IQR 0.5-9.1.

Discussion

Summary of findings

Of 638 included RCTs, about two-thirds mentioned in their results section at least one of the pre-defined categories of participants that might have missing data. The median percentage of participants who were explicitly not followed-up was 5.8% (IQR 2.2-14.8%). When one also includes participants with unclear follow-up status, the total value rises to 11.7% (IQR 5.6-23.7%).

When authors explicitly reported not following-up participants, 54% explicitly reported conducting a complete case analysis; almost all the remainder did not specify how they handled missing data in their analysis. Most RCTs reported neither on missing data separately for different outcomes nor addressed risk of bias associated with missing data. Very few RCTs described a mechanism of missingness (e.g., missing completely at random, missing not at random), and none of the 13 RCTs that imputed outcome took into account the uncertainty associated with imputing outcomes.

Strengths and limitations

As stated in the protocol, this study focused on dichotomous outcome data, given the methodological and statistical issues vary substantively for continuous data [8]. We have excluded time-to-event outcomes for the same reason. We have assessed the reporting and methods for handling missing participant data for continuous outcomes elsewhere [15].

To our knowledge, this is the largest methodological survey on missing data in RCTs and the first to specifically explore how RCT authors report on categories of participants that might have missing data. In addition, our sample of RCTs was not restricted to a specific health-related discipline, which increases the generalizability of our findings.

We used systematic and transparent methods, pilot tested our data abstraction form, and conducted training and calibration exercises for review team members. The core team met on a weekly basis to resolve outstanding issues and conducted triplicate and independent data abstraction as needed to ensure the quality of the data.

Interpretation of findings

Although the median percentage of participants who were explicitly not followed-up was 5.8%, when adding those with unclear follow-up status, the total median percentage doubled to 11.7%. Although less extreme, our results are consistent with those of Marciniak et al., who reported that FDA-calculated loss to follow-up rates were consistently higher than the published rates (median 13% vs. 0.9%, respectively) [6]. Another way of looking at the data is to compare the median percentage of participants who were explicitly not followed-up (5.8%) with the median percentage of participants with unclear follow-up status (9.4%).

An important finding of this survey is that almost one third of the RCTs did not mention any category of participants that might have with missing data. These RCTs either did not have missing data or failed to report the missing data they had. To the extent the latter is the case, results from our sample of RCTs that reported at least one category of participants that might have missing data (n = 400) may be conservative that is underestimating the real extent of missing data. Another important finding is the high

percentage of RCTs not explicitly reporting on the follow-up status of participants (range 32-86%; Table 2). This poor reporting might explain, at least in part, why the majority of SRs fail to adequately report on and handle missing data [7, 12, 13, 16-26].

A rough method to assess the extent of bias associated with missing data is to compare its rate with the risk difference for the outcome of interest. In their previously mentioned study, Marciniak et al. found that the risk difference in the included RCTs had a median of 1.0% (range, 0.2-3.0%) [6]. On the other hand, they found that the median of rates of missing data was 0.4% based on published information, and 13% based on information submitted for FDA review. Although the first median (0.4%) might suggest that the risk of bias associated with missing data is low, the second median (13%) suggests that the risk is actually high. The limitation of their analysis is the lack of comparison of the median missing data rate across study to the median risk difference across studies.

In our study, we compared the risk difference to the missing data rate for each RCT by calculating their ratio. The ratio for participants that were either explicitly not followed-up or with unclear follow-up status had a median of 1.7, which implies that 50% of the RCTs have a ratio of 1.7 or larger. We interpret this that at least half of the RCTs have a high risk of bias associated with missing data. The ratio for participants that were explicitly not followed-up to the risk difference had a 75^{th} percentile of 3.0, which implies that 25% of the RCTs have a ratio of 3 or larger. We interpret this as a substantive minority of RCTs having a high risk of bias associated with missing data. These findings are consistent with those of Marciniak et al. [6].

Additional evidence for the extent to which the missing data can affect the risk of bias comes from a previous study we conducted on the effect of missing data on RCT results [1]. We found that up to one-third of RCTs published in five top general medical journals lose significance when applying plausible assumptions about their loss to follow-up. In that study, rate of participants with missing data was 6%, implying that a much higher percentage of RCTs would likely lose significance when considering the 11.7% of RCT participants that were either explicitly not followed-up or with unclear follow-up status.

We found that complete case analysis was the most frequently reported method for handling missing data in the included RCTs. Although methodologist agrees on the use of complete case analysis to handle missing data in SRs [27], this is not the case when it comes to RCTs. Some methodologists consider that using complete case analysis may bias the estimate of RCT treatment effect [28], whereas others recommend this approach only when data are missing completely at random [29]. A recent simulation study showed that multiple imputations render less biased estimates than other methods including complete case analysis and worst-case scenario when analyzing binary alcohol clinical trial outcomes [30].

Many experts recommend dealing with missing data in RCTs taking into account the mechanism of missingness (e.g., missing completely at random and missing at random) [7, 31-47]. We found that only three of 400 RCTs (<1%) reported on missingness mechanism. One potential explanation is that it might be challenging for RCT authors to judge the mechanism of missingness, which can vary across and even within categories of patients. Consequently, it would not be feasible for SR authors to make imputations taking into account the mechanism of missingness.

Comparison with similar studies

Box 2 compares our findings to those of five recent methods surveys with similar objectives [1, 11-14]. The percentage of RCTs with missing data varied across these six surveys, ranging from 63% to 95%. Similarly, the average percentage of participants with missing data per RCT ranged from 6% to 19%. One reason for these variabilities is the use of different sampling frames across surveys. For example, the survey of RCTs assessing missing quality of life data had the highest percentage of RCTs [12] and that conducted in the context of palliative care had the highest median of participants [11]. Another reason is the use of different definitions of missing data. For example, certain surveys accepted the definition of RCTs being assessed, whereas others used pre-defined categories of participants with missing data. Similarly, the percentage of RCTs reporting on the reasons for missingness varied across surveys. That is likely because of the differential consideration of certain categories as reasons for missingness (e.g., our survey, unlike others, we did not consider 'dead' as equivalent to missing data). Only our survey assessed whether the RCTs reported on missing data per outcome.

The approach to handling missing data was relatively consistent across the surveys, with most RCTs implementing complete case analysis (35-55%). The percentage of RCTs reporting sensitivity analysis to assess the impact of missing data ranged from 5% to 35%. Only our survey assessed whether the RCTs took uncertainty into account when imputing outcomes, as suggested by a number of experts [29, 44, 48, 49]. We found that only 13 of 400 RCTs that included participants with possible missing data reported imputing data. None of these 13 RCTs reported taking uncertainty into account.

Implications of findings

Existing guidance recommends that RCTs report the proportion, reasons, and mechanisms of missing data and how RCT authors accounted for them in the analysis and assessed the associated risk of bias [4, 18, 50]. These recommendations should be adopted by RCT authors and better implemented by journal editors. We suggest that RCT authors additionally report the number of participants with missing data for each outcome and by study arm [3].

For handling missing data in the main analysis, consistent with the suggestion by White et al., we recommend that RCT authors apply plausible assumptions about the outcomes of participants with missing data [51]. These assumptions might depend on (1) the question being examined by the RCT [1], (2) the population involved, (3) the nature of the intervention, and (4) the reason for missingness. Randomized controlled trial authors might need to make different assumptions for different categories of missing data in the same RCT [14, 37]. In terms of assessing the risk of bias associated with missing data, we recommend performing sensitivity analyses to explore the robustness of the results based on the assumption made in the main analysis [51].

There is increasing guidance available to SR authors on handling missing data when synthesizing RCT results [27, 37, 41-43, 45, 47, 52-54]. One proposal is to provide individual participant data or data sharing. The World Health Organization and the International Committee of Medical Journal Editors have highlighted the importance of sharing RCT data [55]. Its significance is related to universal prospective registration and public disclosure of results from all RCTs [55]. Unfortunately, a recent survey of RCTs

published in the British Medical Journal and PLOS Medicine after the adoption of data sharing policies by these journals found that 17 out of 37 RCTs (46%) met criteria for data availability [56]. Both RCT authors and journals editors need to better address missing data in trials reports. Until reporting of missing data is more explicit and transparent, users of the medical literature should take into account potentially missing data in addition to definitely missing data.

Figures, tables, and supplementary data

Legends

Figure 1: The 19 final categories of participants that might have missing data, and the reported follow-up status.

Supplementary file: Numerical example on calculating the ratio of the percentage of participants with missing data to the difference in the event rate

Table 1: General characteristics of the included randomized controlled trials (n=638)

Figure 2: Explanation of information reported in tables 2 and 3 based on an example of reporting on the category 'withdrew consent' by the 638 RCTs

Table 2: Percentage of randomized controlled trials reporting on our pre-defined categories of participants that might have missing data (n= 638 trials)

Table 3: Distribution of percentages of participants across the randomized clinical trials belonging to each category according to follow-up status

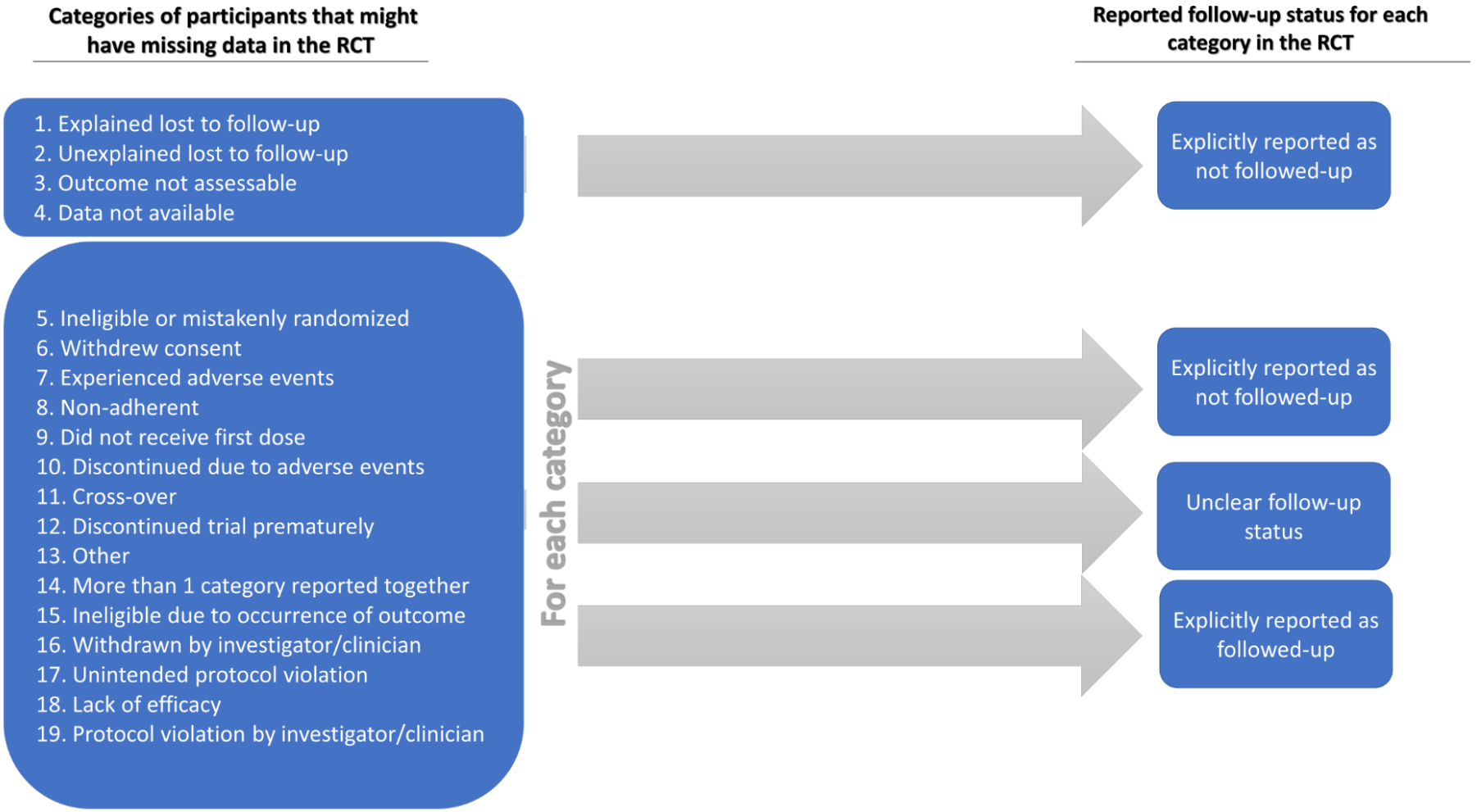
Table 4: Handling in the primary analysis of the pre-defined categories of participants who were explicitly not followed-up or with unclear follow-up status; each RCT could have mentioned more than one category, resulting in a total of 998 instances (n=400 randomized controlled trials that reported at least one category of missing data in the results section)

Table 5: Reporting on, handling of, and assessing risk of bias associated with missing data as defined by trial authors (n=400 randomized controlled trials that reported at least one category of missing data in the results section)

Box 1: Examples of challenges met during data abstraction and management, and the solutions adopted to address them

Box 2: Comparison of the findings of the most recent methods surveys assessing the reporting of missing data, its handling, and the assessment of the associated risk of bias

Figure 1: The 19 final categories of participants that might have missing data, and the reported follow-up status



Supplementary file: Numerical example on calculating the ratio of the percentage of participants with missing data to the difference in the event rate

Randomized controlled trial	**Risk in intervention arm (i):** Events/randomized in intervention arm (%) (n1/N1) %	**Risk in control arm (c):** Events/randomized in control arm (%) (n2/N2) %	**Risk difference (RD):** |Risk in intervention arm - risk in control arm| (%) |i – c|%	**Rate of missing data (MD):** Participants with missing data in both arms/total randomized (%) (m1+m2)/(N1 +N2) %	**Ratio:** Rate of missing data/ risk difference MD/RD
1	(1277/3711) 34.4%	(1277/3711) 34.4%	(1277/3711) 34.4%	(1277/3711) 34.4%	(1277/3711) 34.4%
2	(208/394) 53.6%	(208/394) 53.6%	(208/394) 53.6%	(208/394) 53.6%	(208/394) 53.6%

n1= number of events in the intervention arm
N1- number of participants randomized to the intervention arm
n2= number of events in the control arm
N2- number of participants control to the intervention arm
m1= number of participants with missing data in the intervention arm
m2= number of participants with missing data in the control arm

Table 1: General characteristics of the included randomized controlled trials (n=638)

Variable	N (%)
Type of paper	
Full text article	617 (96.7)
Abstract/research letter	21 (3.3)
Year of publication (median [IQR])	2005 (1998 – 2008)
Language of report	
English	606(95)
Non-English	32(5)
Source of funding of trial	
Private for profit	220 (34.5)
Private not for profit	118 (18.5)
Government	107 (16.8)
Not funded	16 (2.5)
Not reported	284 (44.5)
Type of intervention	
Pharmacological	382 (58.5)
Surgery/invasive procedure	169 (25.9)
Other	102 (15.6)
Planned follow-up time reported	509 (79.8)
Planned follow-up time in months [median (IQR)] [a]	6 (1 - 12)
Time-to-occurrence reported	95 (14.9)
Time-to-occurrence in months [median (IQR)] [b]	1 (0.5 - 6)
Number randomized, intervention group [median (IQR)]	69 (31 – 167)
Number randomized, control group [median (IQR)]	66 (32 - 162)
Reported type(s) of analyses	
Intention to treat (ITT)	233(36.5)
Modified ITT [c]	21(3.3)
Per protocol	43(6.7)
As treated	12(1.9)
None of the above	388(60.8)

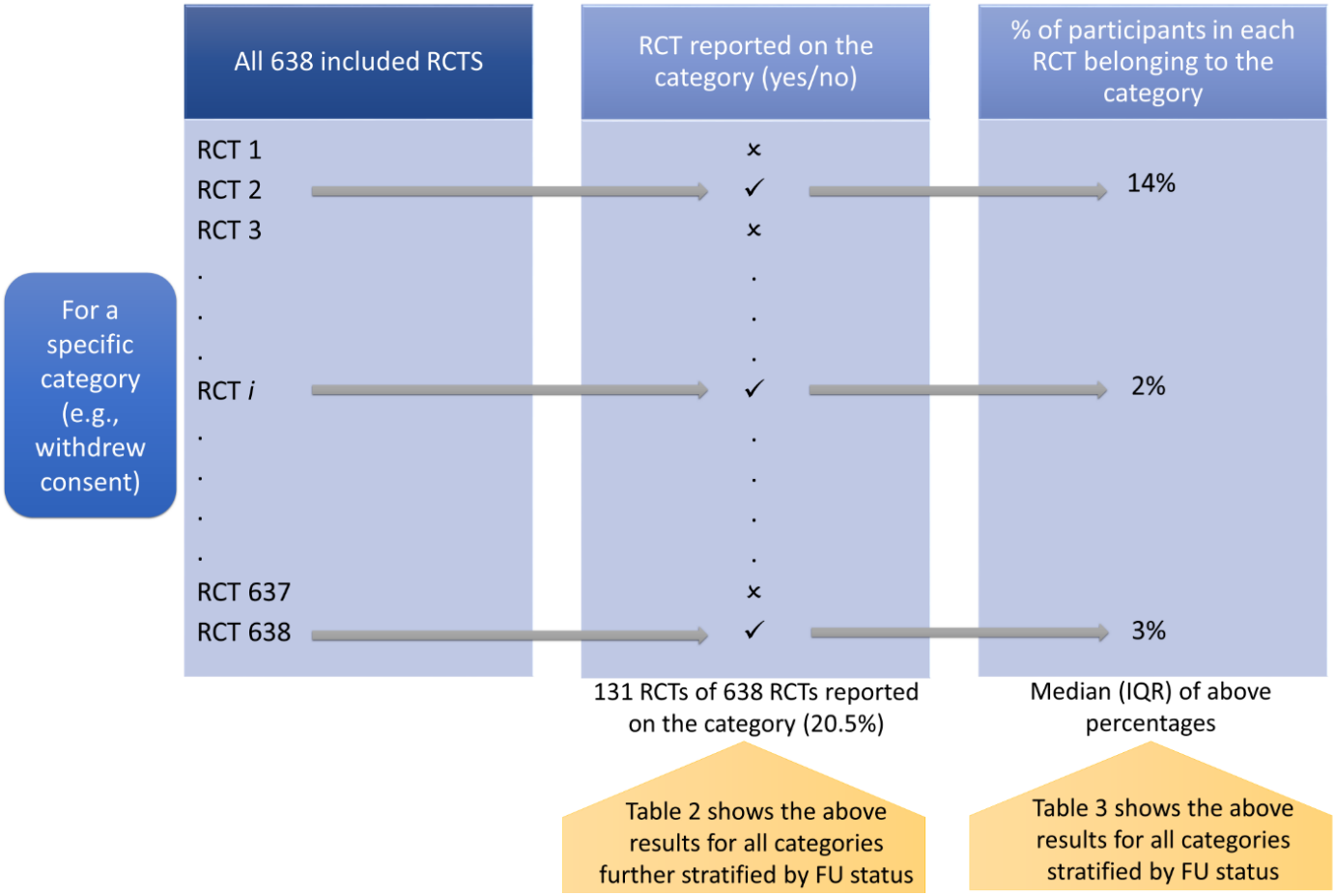
Abbreviations: **ITT**: intention-to-treat; **IQR**: interquartile range.

[a] N= 509, number of trials reporting planned follow-up time

[b] N= 95, number of trials reporting time to occurrence

[c] We reported on whether the authors reported on using 'modified ITT', either by using the terminology explicitly, or by describing a modification to the intention-to-treat analysis (i.e., analyzed all participants as they were randomized with the exception of certain category of participants they decided to exclude).

Figure 2: Explanation of information reported in tables 2 and 3 based on an example of reporting on the category 'withdrew consent' by the 638 RCTs



Abbreviations: **FU**: follow-up; **IQR**: Interquartile range; **RCT**: randomized controlled trial

Table 2: Percentage of randomized controlled trials reporting on our pre-defined categories of participants that might have missing data (n= 638 trials)

Categories	RCTs reporting on the category	RCTs reporting on follow-up status for each category [a]		
	N (%)	Explicitly FU	Explicitly not FU	Unclear FU status
Unexplained lost to follow-up	161 (25.2)	0	161(100.0)	0
Ineligible or mistakenly randomized	142 (22.3)	10 (7.0)	36 (25.4)	96(67.6)
Experienced adverse events	135 (21.2)	92 (68.1)	0	43(31.9)
Withdrew consent	131 (20.5)	6 (4.6)	23 (17.6)	102(77.9)
Non-adherent	131 (20.5)	14 (10.7)	15 (11.4)	102(77.9)
Did not receive the first dose	89 (13.9)	4 (4.5)	12 (13.5)	73 (82.0)
Discontinued due to adverse events	75 (11.8)	7 (9.3)	23 (30.7)	45(60.0)
Cross-over	62 (9.7)	21 (33.9)	2 (3.2)	39(62.9)
Discontinued trial prematurely	55 (8.6)	8 (14.5)	4 (7.3)	43(78.2)
Other [b]	41 (6.4)	1 (2.4)	5 (12.2)	35(85.4)
Explained lost to follow-up	30 (4.7)	0	30 (100.0)	0
More than one category reported together	28 (4.4)	1(3.6)	6 (21.4)	21(75.0)
Outcome not assessable	18 (2.8)	0	18 (100)	0
Data not available	18 (2.8)	0	18 (100.0)	0
Ineligible due to occurrence of outcome	14 (2.2)	1(7.1)	1 (7.1)	12(85.8)
Withdrawn by investigator/clinician	13 (2.0)	0	4 (30.8)	9(69.2)
Unintended protocol violation	9 (1.4)	1(11.1)	1 (11.1)	7(77.8)
Lack of efficacy	8 (1.3)	0	2 (25.0)	6(75.0)
Protocol violation by investigator/clinician	4 (0.6)	0	1 (25.0)	3(75.0)

Abbreviations: **FU**: followed-up; **RCT**: randomized controlled trial.

[a] The denominator is the number of trials reporting on this category.

[b] Two trials listed a second category as 'other', the percentages of participants were 0.4% and 2.4%. Both trials did not explicitly report on the follow-up status of these participants.

Table 3: Distribution of percentages of participants across the randomized clinical trials belonging to each category according to follow-up status

Categories	Explicitly not FU		Unclear FU status		Explicitly not FU and Unclear FU status	
	N of RCTs	Median % (IQR)	N of RCTs	Median% (IQR)	N of RCTs	Median% (IQR)
Unexplained lost to follow-up	161	4.1 (1.2 – 10.3)	0	-	161	4.1 (1.2 – 10.3)
Ineligible or mistakenly randomized	36	4.1 (1.8 – 10.0)	96	2.3 (1.0 – 6.5)	132	2.9 (1.0 – 7.7)
Experienced adverse events	0	-	43	11.0 (4.4 – 37.2)	43	11.0 (4.4 – 37.2)
Withdrew consent	23	3.7 (1.7 – 5.7)	102	2.6 (1.1 – 6.3)	125	2.9 (1.1 – 6.1)
Non-adherent	15	3.3 (1.7 – 8.2)	102	6.9 (1.8 – 15.1)	117	5.4 (1.8 – 14.5)
Did not receive the first dose	12	1.0 (0.5 – 1.7)	73	1.6 (0.9 – 3.6)	85	1.3 (0.9 – 3.1)
Discontinued due to adverse events	23	3.7 (2.2 – 6.5)	45	4.0 (2.1 – 6.1)	68	4.0 (2.2 – 6.3)
Cross-over	2	4.3, 11.0	39	3.0 (1.2 – 5.0)	41	3.3 (1.3 – 5.3)
Discontinued trial prematurely	4	4.9 (0.9 – 10.7)	43	6.3 (3.5 – 18.1)	47	6.3 (3.2 – 16.4)
Other	5	1.5 (0.9 – 2.4)	35	2.1 (1.0 – 4.1)	40	2.1 (1.0 – 4.0)
Explained lost to follow-up	30	3.6 (2.0 – 7.3)	0	-	30	3.6 (2.0 – 7.3)
More than one category reported together	6	10.7 (3.9 – 18.7)	21	5.4 (3.5 – 9.2)	27	6.5 (4.1 – 13.4)
Outcome not assessable	18	27.3 (3.7 – 28.7)	0	-	18	27.3 (3.7 – 28.7)
Data not available	18	2.6 (0.9 – 5.3)	0	-	18	2.6 (0.9 – 5.3)
Ineligible due to occurrence of outcome	1	4.3	12	1.9 (0.3 – 4.1)	13	2.4 (0.4 – 4.2)
Withdrawn by investigator/clinician	4	1.5 (0.9 – 3.0)	9	0.9 (0.4 – 5.8)	13	1.3 (0.6 – 4.2)
Unintended protocol violation	1	0.4	7	1.4 (0.3 – 4.9)	8	0.9 (0.3 – 4.9)
Lack of efficacy	2	0.1, 7.9	6	2.5 (1.2 – 6.2)	8	2.5 (0.6 – 7.9)
Protocol violation by investigator/clinician	1	0.3%	3	0.6%, 0.7%, 1.3%	4	0.7 (0.4 – 1.2)
All categories combined	**256**	**5.8 (2.2 - 14.8)**	**288**	**9.7 (4.1 – 19.9)**	**400**	**11.7 (5.6 – 23.7)**

Abbreviations: **FU**: followed-up; **IQR**: Interquartile range; **RCT**: randomized controlled trial

Table 4: Handling in the primary analysis of the pre-defined categories of participants who were explicitly not followed-up or with unclear follow-up status; each RCT could have mentioned more than one category, resulting in a total of 998 instances (n=400 randomized controlled trials that reported at least one category of missing data in the results section)

Categories	Dealing method in the analysis							
	Explicitly not followed-up				Unclear follow-up status			
	n	CCA	Other method*	NR	n	CCA	Other method*	NR
Unexplained loss to FU	161	74 (46.0)	5 (3.1)	82 (50.9)	0	0	0	0
Ineligible or mistakenly randomized	36	33 (91.7)	0	3 (8.3)	96	47 (49.0)	0	49 (51.0)
Experienced adverse events	0	0	0	0	43	0	0	43 (100.0)
Withdrew consent	23	13 (56.5)	1 (4.3)	9 (39.2)	102	31 (30.4)	2(2.0)	69 (67.6)
Non-adherent	15	11 (73.3)	1 (6.7)	3 (20.0)	102	25 (24.5)	1(1.0)	76 (74.5)
Did not receive the first dose	12	11 (91.7)	0	1 (8.3)	73	37 (51.7)	0	36 (49.3)
Discontinued due to adverse events	23	8 (34.8)	1 (4.3)	14 (60.9)	45	12 (26.7)	0	33 (73.3)
Cross-over	2	1 (50.0)	0	1 (50.0)	39	2 (5.1)	1(2.6)	36 (92.3)
Discontinued trial prematurely	4	3 (75.0)	0	1 (25.0)	43	7 (16.3)	1(2.3)	35 (81.4)
Other	5	2 (40.0)	1 (20.0)	2 (40.0)	35	5 (14.3)	0	30 (85.7)
Explained loss to FU	30	9 (30.0)	0	21 (70.0)	0	0	0	0
More than one category reported	6	2 (33.3)	0	4 (66.7)	21	10 (46.7)	2(9.5)	9 (42.8)
Outcome not assessable	18	11 (61.1)	0	7 (38.9)	0	0	0	0
Data not available	18	13 (72.2)	2 (11.1)	3 (16.7)	0	0	0	0
Ineligible due to occurrence of outcome	1	0	0	1 (100.0)	12	5 (41.7)	0	7 (58.3)
Withdrawn by investigator/clinician	4	2 (50.0)	0	2 (50.0)	9	1 (11.1)	0	8 (88.9)
Unintended protocol violation	1	1 (100.0)	0	0	7	3 (42.9)	0	4 (57.1)
Lack of efficacy	2	0	0	2 (100.0)	6	1 (16.7)	0	5 (83.3)
Protocol violation by investigator/clinician	1	1 (100.0)	0	0	3	1 (33.3)	0	2 (67.7)
Total	**362**	**195 (53.9)**	**11 (3.0)**	**156 (43.1)**	**636**	**187 (29.4)**	**7 (1.1)**	**442 (69.5)**

Abbreviations: **CCA**: complete case analysis; **FU**: follow-up; **NR**: not reported
*Other method: including making assumptions and imputations

Table 5: Reporting on, handling of, and assessing risk of bias associated with missing data as defined by trial authors (n=400 randomized controlled trials that reported at least one category of missing data in the results section)

Variable	N (%)
Missing data explicitly reported in the results section	
Not separate for different outcomes	397(99.3)
For each outcome separately	3(0.7)
Number of participants with missing data reported	
Yes, overall (both arms combined)	105(26.2)
Yes, per arm	352(88.3)
No	7 (1.8)
Baseline characteristics of participants with missing data reported	
Yes, missing data group vs. non-missing data group	7(1.7)
Yes, missing data in intervention arm vs. missing data in control arm	4(1.0)
No	389 (97.3)
Number of participants with missing data compared between the two study arms	24(6.0)
Mechanisms of missingness for missing data described the missing data (e.g., missing completely at random, missing not at random)	3 (0.7)
Uncertainty associated with imputing outcomes taken into account, in case imputation was done	
Imputed outcomes & took uncertainty into account	0
Imputed outcomes but did not take uncertainty account	13(3.2)
Not applicable, no imputation for missing data	387(96.8)
Justification for the method used to handle missing data provided	
Yes	2(0.5)
No, missing data was handled but not justified	190(47.5)
Not applicable, missing data was not handled	208(52.0)
Method reported to be used for judging risk of bias associated with missing data	
Complete Case Analysis	6(1.5)
None had event	5(1.2)
All had event	3(0.7)
Same event rate	0
Worst Case Scenario	5(1.2)
Best Case Scenario	3(0.7)
Other [a]	3(0.7)
Single Imputation	0
Multiple Imputation	0
Mixed Effect model	0
Unclear which method used	4(1.0)
No method reported	378 (94.5)

[a] Other methods included:

- Compared the effectiveness of intervention between completed and censored participants
- Assumed rate of primary efficacy endpoint in sensitivity analysis adjusted for missing values during intended treatment period – primary subjects and non-primary subject
- Treated participants who withdrew, were lost to follow-up, or died as censored data for survival analysis if the event under investigation had not occurred

Box 1: Examples of challenges met during data abstraction and management, and the solutions adopted to address them

Challenge	Solution
• Certain categories of participants that might have missing data reported by RCT authors did not fit in the original list of those categories.	• We refined the original list to accommodate new categories that emerged from data abstraction and did not fit already defined categories.
• Trial author may have counted some participants under more than one of the pre-defined categories of participants that might have missing data, e.g., 'discontinued trial prematurely' and 'non-adherent'. • In some instances, this double counting was obvious (e.g., in the CONSORT flow diagram), in others it was not.	• Whenever double counting was obvious, we listed these participants under only one category.
• Some RCTs did not clearly report the number of events of the completers.	• We used the number of events as reported in the meta-analysis; we implicitly assumed that the SR authors used accurate numbers (e.g., by contacting the RCT authors).
• Few RCTs excluded participants with available outcome data.	• We considered these participants as not having missing data.
• Some RCTs reported on the percentages of participants belonging to the pre-defined categories but not their count. It was not clear whether the denominator was the number of participants randomized, the number of participants who received treatment, or number of participants who were compliant.	• We made our best guess of which denominator the authors used. When that was not possible, we used the number of participants randomized.
• In some instances, the population of interest of the SR was a subgroup population of the included RCT, and the RCT authors did not report on the categories of missing data within that subgroup.	• For each category, we multiplied the number of participants with missing data in the overall study population by the proportion of the participants in the subgroup of interest

Box 2: Comparison of the findings of the most recent methods surveys assessing the reporting of missing data, its handling, and the assessment of the associated risk of bias

Study	Sampling frame; N	Description of missing data	% of RCTs reporting missing data	% of participants with missing data	% of RCTs reporting on reasons for missing data	% of RCTs comparing baseline characteristics between patients with observed and missing	Most common method of handling missing data (%)	% of RCTs reporting on sensitivity analysis to assess the impact of missing data	% of RCTs reporting on missing data per outcome	% of RCTs reporting on mechanism of missing data	% of RCTs taking uncertainty into account when imputing outcomes
Current study	100 SRs with statistically significant effect published in 2012;	19 categories of participants that might have missing data'	63 (400 RCTs)	Median of 11.7% across the 400 RCTs with missing data	100%, as study defined missing data based on reasons	2% of the 400 RCTs with missing data	CCA (55% in the 400 RCTs with missing data)	5% of the 400 RCTs with missing data	3% of the 400 RCTs with missing data	3% of the 400 RCTs with missing data	0% of the 13 RCTs that imputed outcomes for at least one category
Hussain 2017(7)	Palliative care RCTs (January 2009-April 2014); N=108	Missing data'	86 (93 RCTs)	19% among all participants in the 93 RCTs with missing data	71% of the 93 RCTs with missing data	13% of the 93 RCTs with missing data	CCA (37% in the 93 RCTs with missing data)	16% of the 93 RCTs with missing data	Not assessed	3% in the 93 RCTs with missing data	Not assessed
Fielding 2016 (6)	Top 4 medical journals (50% random sample from 2013 -2014 RCTs on	Missing quality of life data	95 (83 RCTs)	Median is between 5 and 10% *	Not assessed	Not assessed	CCA (67% in the 83 RCTs with missing data) +	15% (this was based on the use of imputations)	Not assessed	13% of the 83 RCTs with missing data	Not assessed
Bell 2014 (5)	Top 4 medical journals (July-December 2013 RCTs); N=77	Some missing data	95 (73 RCTs)	Median of 9% in the 73 RCTs with missing data	90% of the 73 RCTs with missing data	12% of the 73 RCTs with missing data	CCA (45% in the 73 RCTs with missing data	35% of the 73 RCTs with missing data	Not assessed	Not assessed	Not assessed

Study	Sampling frame; N	Description of missing data	% of RCTs reporting missing data	% of participants with missing data	% of RCTs reporting on reasons for missing data	% of RCTs comparing baseline characteristics between patients with observed and missing outcomes	Most common method of handling missing data (%)	% of RCTs reporting on sensitivity analysis to assess the impact of missing data	% of RCTs reporting on missing data per outcome	% of RCTs reporting on mechanism of missing data	% of RCTs taking uncertainty into account when imputing outcomes
Powney 2014 (8)	100 RCTs reporting on longitudinal or repeated measurements (2005-2012);	Missing data	91 (91 RCTs)	Not assessed	38% of the 91 RCTs with missing data	Not assessed	CCA (35% in the 91 RCTs)	Not assessed	Not assessed	Not assessed	Not assessed
Akl 2012 (4)	Top 5 medical journals (2008 RCTs reporting statistically significant effects); N=235	Some loss to follow-up	81 (191 RCTs)	Median of 6% across the 191 RCTs with missing data	Not assessed	Not assessed	CCA (23% in the 191 RCTs with missing data)	Not assessed	Not assessed	Not assessed	Not assessed

Abbreviations: **CCA:** complete case analysis; **missing data:** missing data; **RCTs**: randomized controlled trials; **SRs**: systematic reviews
∗ Excluding the studies with no missing data (5%) and studies where the proportion of missing data is unclear (35%).
+The primary analysis strategy for the quality of life outcome was a complete case analysis for 42/87 RCTs (48%).

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. Rombach I, Rivero-Arias O, Gray AM, Jenkinson C, Burke O. The current practice of handling and reporting missing outcome data in eight widely used PROMs in RCT publications: a review of the current literature. Qual Life Res 2016; 25:1613-23.
3. Akl EA, Kahale LA, Ebrahim S, Alonso-Coello P, Schunemann HJ, Guyatt GH. Three challenges described for identifying participants with missing data in trials reports, and potential solutions suggested to systematic reviewers. J Clin Epidemiol 2016; 76:147-54.
4. National Research Council. The prevention and treatment of missing data in clinical trials. Washington, DC: National Academies Press; 2010.
5. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al. The prevention and treatment of missing data in clinical trials. N Engl J Med 2012; 367:1355-60.
6. Marciniak TA, Cherepanov V, Golukhova E, Kim MH, Serebruany V. Drug discontinuation and follow-up rates in oral antithrombotic trials. JAMA Intern Med 2016; 176(2):257-9.
7. Gewandter JS, McDermott MP, McKeown A, Smith SM, Williams MR, Hunsinger M, et al. Reporting of missing data and methods used to accommodate them in recent analgesic clinical trials: ACTTION systematic review and recommendations. Pain 2014; 155(9):1871-7.
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. Kahale LA, Diab B, Brignardello-Petersen R, Agarwal A, Mustafa RA, Kwong J, et al. Systematic reviews do not adequately report, or address missing outcome data in their analyses: a methodological survey. J Clin Epidemiol 2018; 99:14-23.
10. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)–a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009; 42(2):377-81.
11. Hussain JA, Bland M, Langan D, Johnson MJ, Currow DC, White IR. Quality of missing data reporting and handling in palliative care trials demonstrates that further development of the CONSORT statement is required: a systematic review. J Clin Epidemiol 2017; 88:81-91.
12. Fielding S, Ogbuagu A, Sivasubramaniam S, MacLennan G, Ramsay CR. Reporting and dealing with missing quality of life data in RCTs: has the picture changed in the last decade? Qual Life Res 2016; 25:2977-83.
13. Bell ML, Fiero M, Horton NJ, Hsu CH. Handling missing data in RCTs; a review of the top medical journals. BMC Med Res Methodol 2014; 14:118.
14. Powney M, Williamson P, Kirkham J, Kolamunnage-Dona R. A review of the handling of missing longitudinal outcome data in clinical trials. Trials 2014; 15(1):237.
15. Zhang Y, Florez ID, Colunga Lozano LE, Aloweni FAB, Kennedy SA, Li A, et al. A systematic survey on reporting and methods for handling missing participant data for

continuous outcomes in randomized controlled trials. J Clin Epidemiol 2017; 88:57-66.
16. Hussain JA, White IR, Langan D, Johnson MJ, Currow DC, Torgerson D, et al. Missing data in randomised controlled trials evaluating palliative interventions: a systematic review and meta-analysis. Lancet 2016;387:S53.
17. Adewuyi TE, MacLennan G, Cook JA. Non-compliance with randomized allocation and missing outcome data in randomized controlled trials evaluating surgical interventions: a systematic review. BMC Res Notes 2015; 8:403.
18. Akl EA, Shawwa K, Kahale LA, Agoritsas T, Brignardello-Petersen R, Busse JW, et al. Reporting missing participant data in randomised trials: systematic survey of the methodological literature and a proposed guide. BMJ Open 2015; 5(12):e008431.
19. Fielding S, Maclennan G, Cook JA, Ramsay CR. A review of RCTs in four medical journals to assess the use of imputation to overcome missing data in quality of life outcomes. Trials 2008; 9:51.
20. Fiero MH, Huang S, Oren E, Bell ML. Statistical analysis and handling of missing data in cluster randomized trials: a systematic review. Trials 2016; 17:72.
21. Karlson CW, Rapoff MA. Attrition in randomized controlled trials for pediatric chronic conditions. J Pediatr Psychol 2009; 34(7): 782-93.
22. Masconi KL, Matsha TE, Echouffo-Tcheugui JB, Erasmus RT, Kengne AP. Reporting and handling of missing data in predictive research for prevalent undiagnosed type 2 diabetes mellitus: a systematic review. EPMA J 2015; 6(1):7.
23. Spineli LM. Missing binary data extraction challenges from Cochrane reviews in mental health and Campbell reviews with implications for empirical research. Res Synth Methods 2017; 8(4):514-25.
24. Wahlbeck K, Tuunainen A, Ahokas A, Leucht S. Dropout rates in randomised antipsychotic drug trials. Psychopharmacology 2001; 155(3):230-3.
25. Wood AM, White IR, Thompson SG. Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. Clin Trials 2004; 1:368-76.
26. Ibrahim F, Tom BD, Scott DL, Prevost AT. A systematic review of randomized controlled trials in rheumatoid arthritis: the reporting and handling of missing data in composite outcomes. Trials 2016; 17(1):272.
27. 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.
28. Groenwold RH, Moons KG, Vandenbroucke JP. Randomized trials with missing outcome data: how to analyze and what to report. CMAJ 2014; 186(15):1153-7.
29. Little RJ, Rubin DBIn: Statistical analysis with missing data, 333. New York: John Wiley & Sons; 2014.
30. Hallgren KA, Witkiewitz K, Kranzler HR, Falk DE, Litten RZ, O'malley SS, et al. Missing data in alcohol clinical trials with binary outcomes. Alcohol Clin Exp Res 2016; 40 (7): 1548-57.
31. Barnes SA, Mallinckrodt CH, Lindborg SR, Carter MK. The impact of missing data and how it is handled on the rate of false-positive results in drug development. Pharm Stat 2008; 7(3):215-25.
32. Bell ML, Fairclough DL. Practical and statistical issues in missing data for longitudinal patient-reported outcomes. Stat Methods Med Res 2014; 23(5):440-59.

33. Bell ML, Kenward MG, Fairclough DL, Horton NJ. Differential dropout and bias in randomised controlled trials: when it matters and when it may not. BMJ 2013; 346:e8668.
34. Buehl W, Heinzl H, Mittlboeck M, Findl O. Statistical problems caused by missing data resulting from neodymium:YAG laser capsulotomies in long-term posterior capsule opacification studies: problem identification and possible solutions. J Cataract Refract Surg 2008; 34:268-73.
35. Fielding S, Fayers P, Ramsay C. Predicting missing quality of life data that were later recovered: an empirical comparison of approaches. Clin Trials 2010; 7:333-42.
36. Fielding S, Fayers PM, Ramsay CR. Investigating the missing data mechanism in quality of life outcomes: a comparison of approaches. Health Qual Life Outcomes 2009;7:57.
37. Higgins JP, White IR, Wood AM. Imputation methods for missing outcome data in meta-analysis of clinical trials. Clin Trials 2008; 5:225-39.
38. Ibrahim JG, Chu H, Chen MH. Missing data in clinical studies: issues and methods. J Clin Oncol 2012; 30:3297-303.
39. Jorgensen AW, Lundstrom LH, Wetterslev J, Astrup A, Gotzsche PC. Comparison of results from different imputation techniques for missing data from an anti-obesity drug trial. PLoS One 2014; 9:e111964.
40. Liu M, Wei L, Zhang J. Review of guidelines and literature for handling missing data in longitudinal clinical trials with a case study. Pharm Stat 2006; 5 (1):7-18.
41. 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.
42. 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.
43. Spineli LM, Higgins JP, Cipriani A, Leucht S, Salanti G. Evaluating the impact of imputations for missing participant outcome data in a network meta-analysis. Clin Trials 2013; 10:378-88.
44. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009; 338:b2393.
45. 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.
46. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. BMJ 2011; 342:d40.
47. White IR, Welton NJ, Wood AM, Ades AE, Higgins JP. Allowing for uncertainty due to missing data in meta-analysis–part 2: hierarchical models. Stat Med 2008; 27:728-45.
48. Harel O, Mitchell EM, Perkins NJ, Cole SR, Tchetgen EJ, Sun B, et al. Multiple imputation for incomplete data in epidemiologic studies. Am J Epidemiol 2018; 187:576-84.
49. Rubin DBIn: Multiple imputation for nonresponse in surveys, 81. New York: John Wiley & Sons; 2004.
50. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux P, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340:c869.
51. White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. Clin Trials 2012; 9:396-407.

52. Akl EA, Johnston BC, Alonso-Coello P, Neumann I, Ebrahim S, Briel M, et al. Addressing dichotomous data for participants excluded from trial analysis: a guide for systematic reviewers. PLoS One 2013; 8:e57132.
53. Higgins J, Green S. Cochrane Handbook for systematic reviews of interventions version 5.1.0 2011: The Cochrane Collaboration. Available at www.handbook.cochrane.org.
54. 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.
55. Taichman DB, Sahni P, Pinborg A, Peiperl L, Laine C, James A, et al. Data sharing statements for clinical trials: a requirement of the International Committee of Medical Journal. Ann Intern Med 2017; 167:63-5.
56. Naudet F, Sakarovitch C, Janiaud P, Cristea I, Fanelli D, Moher D, et al. Data sharing and reanalysis of randomized controlled trials in leading biomedical journals with a full data sharing policy: survey of studies published in the BMJ and PLOS Medicine. BMJ 2018;360:k400

Chapter 4

A guidance was developed to identify participants with missing outcome data in randomized controlled trials

Kahale LA, Guyatt GH, Agoritsas T, Briel M, Busse JW, Carrasco-Labra A, Khamis AM, Zhang Y, Hooft L, Scholten RJ, Akl EA, A guidance was developed to identify participants with missing outcome data in randomized controlled trials, Journal of Clinical Epidemiology (2019), doi:https://doi.org/10.1016/j.jclinepi.2019.07.003

Abstract

Background:

In order for authors of systematic reviews to address missing data in randomized controlled trials (RCTs), they need to first identify the number of trial participants with missing data. Our objective is to provide guidance for authors of systematic reviews on how to identify participants with missing outcome data in reports of RCTs.

Methods:

Guidance statements were informed by a review of studies addressing the topic of missing data and an iterative process of feedback and refinement, through meetings involving experts in health research methodology and authors of systematic reviews.

Results:

The proposed guidance includes: (1) definitions of key terms, (2) 19 categories of participants described in RCT reports and who might have missing data, and (3) a flowchart on how to judge the outcome data missingness for each category. The judgment of missingness relies on how trial authors report on the categories and handle them in their analyses. Practically, for their primary analysis, systematic reviewer authors should choose how to identify participants with missing outcome data (i.e., use either 'definitely missing data' or 'total possible missing data'), then select a method for handling missing data in meta-analysis. Sensitivity analyses should be undertaken to explore consistency with competing options for classifying patients as having missing data.

Conclusion:

Adopting the proposed guidance will help promote transparency and consistency regarding how missing data is managed in systematic reviews.

Introduction

Authors of systematic reviews are frequently confronted with missing data for on one or more outcomes of trial participants. Recent studies found that 42% of systematic reviews (1) and 63% of randomized controlled trials (RCTs) reported on participants with missing data (2). Missing data may bias results of RCTs when outcomes of those missing differ systematically from those who have been followed-up. Thus, inferences from reviews of RCTs may be misleading if trial authors do not handle missing data appropriately (3-5).

Guidance for authors of systematic reviews on how to assess the risk of bias associated with missing data in a meta-analysis and how to handle this is available for both dichotomous and continuous outcomes (6-14). However, in order to follow this guidance, they first need to identify for every outcome how many trial participants actually have missing data. This task can become quite complex, if RCTs do not clearly report this information. Identifying missing outcome data in trial reports is associated with three main challenges (see box 1) (3).

Box 1: Three main challenges faced by authors of systematic reviews when identifying missing outcome data (3)

1. Although systematic reviewers require information about missing data to be reported by outcome, trialists typically report the information by participant;
2. It is not always clear whether trialists have successfully followed participants in certain categories (e.g., those who withdrew consent); that is, whether some categories of participants did or did not have missing data;
3. It is not always clear how the trialists dealt with missing data in their analyses.

As an example of unclear reporting, RCTs do not always specify whether participants categorized as 'lost to follow-up' actually developed an event for the outcome of interest before they were lost to follow-up. Although systematic reviewers typically consider

participants lost to follow-up as having missing data for all outcomes, this may not be the case (3). Ideally, RCTs should report missing data by outcome, but very few RCTs do so; a methodologic survey of 638 RCTs found that 0.7% reported missing data by outcome (2). Many RCTs do not report the method(s) used for handling missing data (43% as a conservative estimate) (2).

Some trialists might assume that participants with missing data experienced an event for the outcome of interest and include them in the numerator (number of events for that outcome), but without specifying this approach. This is a standard approach in tobacco cessation trials. According to Russell standards, participants who are analyzed are counted as smokers if their smoking status at final follow-up cannot be determined (15, 16). Indeed, Foulds et al telephoned participants who were lost to follow-up in a hospital-based smoking cessation trial and reported that 100% of these participants relapsed (17, 18). Similarly, in mental health research, many experts observed that missing data are likely not to be missing at random (i.e., the probability that an observation is missing depends on the unseen observations themselves); assuming that they developed an event for the outcome of interest is very plausible (19, 20). If trialists have already included events for some of those they report to have missing data, and review authors assume that participants with missing data had an event for the outcome of interest, the result will be double counting (3).

The extent of missing data in RCTs might be larger than what authors explicitly report. For example, one study compared 'loss to follow-up rates' in published reports of RCTs of oral antithrombotic agents with loss to follow-up rates derived from more detailed documents made available to the FDA for the same RCTs (21). They found a large discrepancy between the median published rate of all 'missing follow-up categories' (0.9%), and the median of the FDA-calculated loss to follow-up rates (13%). We found similar results in a recently published review of 638 RCTs; specifically, the median percentage of participants who were explicitly not followed-up was 5.8%, but this value increased to 11.7% when considering patients for whom the follow-up status was unclear (2).

Given the above challenges, and until reporting of missing data becomes more explicit, authors of systematic reviews require guidance on how to identify participants with missing outcome data in RCT reports. We are not aware of any such existing guidance. In particular, the recent update of the Cochrane handbook does not address the issue (14).

Objectives

The objective of this paper is to provide guidance for authors of systematic reviews on how to identify participants with missing outcome data in reports of RCTs.

Methods

To inform our guidance, we reviewed the following studies on the topic of missing data:

- Proposed approaches for reporting and handling missing data in RCTs (22) and in systematic reviews (23, 24);
- Conceptual paper on challenges faced by systematic reviewers while identifying trial participants with missing data (3);
- Methodological surveys on the reporting and handling of missing data in RCTs (25) and systematic reviews (1, 26);
- Impact of missing data on effect estimates in RCTs (27);
- Impact of missing data on effect estimates in systematic reviews (unpublished data);
- Guidance for handling missing data of dichotomous (6) and continuous outcomes in systematic reviews (7, 28);
- GRADE guidance for assessing risk of bias associated with missing data in a body of evidence (8).

Accordingly, we developed a draft guidance for authors of systematic reviews on how to identify participants with missing outcome data in reports of RCTs. Then, we revised guidance using an iterative process of feedback and refinement through three face-to-face meetings and several teleconferences involving experts in health research methodology, and authors of systematic reviews as the end users. In addition, we conducted two workshops at Cochrane colloquia for further feedback.

This guidance is for meta-analyses of group-level data from RCTs. It does not address methods for meta-analyses of individual participant data.

Results

We describe below the proposed guidance for identifying missing outcome data which includes: (1) the definitions of key terms, (2) the categories of participants described in RCT reports that might be associated with missing data, and (3) how to judge the outcome data missingness for these categories.

Definitions

We used the following definitions (see below figure 1):

- Missing data: outcome data from included RCTs that are not available to the authors of systematic reviews, whether from published RCT reports or through contact with trialists.
- Definitely missing data: participants clearly have missing outcome data based on the RCT reporting.
- Potentially missing data: participants potentially have missing outcome data, but it is not explicitly reported in the RCT report.
- Total possible missing data: participants who either have definitely or potentially missing outcome data.

Categories

We developed a draft taxonomy of categories of participants described in RCT reports and who might have missing data. We refined this taxonomy iteratively through its application to data extraction in a succession of published studies: 235 RCTs (27), 202 systematic reviews (6) , 100 systematic reviews (1), 200 RCTs (25), and 638 RCTs (2) (for a total of 1,073 RCTs and 302 systematic reviews). We labeled categories to reflect wording used in RCT reports, i.e., the presentation that systematic review authors actually face: we are not suggesting using these categories when reporting RCTs. Table 1 lists the 19 final categories of participants that might have missing data along with their description.

We considered two additional categories reported in trials ('dead' and 'excluded as part of center exclusion'), but decided not to consider them as missing data as explained here:

- 'Dead' category: As 'death' is a competing outcome, we consider that participants described as 'dead' to not have missing data. In other words, the interpretation of the outcome of interest should consider these participants as 'dead', and not as having missing data. See box 2 for example.

Box 2: Death as a competing outcome

> A competing risk is an event that either prevents the observation of an event of interest or modifies the chance that this event occurs (57). Consider an RCT comparing a palliative care intervention to standard of care, showing an increased incidence of death but improved quality of life among those who survive. When analyzing the results for the quality of life outcome, one should not impute data for those who died before their quality of life could be assessed at the end of the RCT.

- 'Excluded as part of center exclusion' category: In multicenter trials, individual centers may be excluded from the study due to a specific reason (e.g., low recruitment, non-adherence to trial protocol). Participants, who have been already enrolled by those centers prior to the decision of exclusion, will be excluded from the study and not followed-up. We consider participants belonging to this category to not have missing data and to be appropriately excluded from the denominator of trial participants.

Judging outcome data missingness

Figure 2 shows the flowchart illustrating our proposed guidance on how authors of systematic reviews could judge the missingness of outcome data in RCT reports.

- 'Dead': As noted earlier, we consider participants belonging to this category to definitely not have missing data;
- 'Excluded as part of center exclusion': As noted earlier, we consider participants belonging to this category to definitely not have missing data;

- For the first four categories in Table 1 ('explained lost to follow-up', unexplained lost to follow-up', 'outcome not assessable', or 'data not available'), we consider participants belonging to these categories to definitely have missing data;
- For the remaining 15 categories, the guidance for the remaining categories is based on the reporting of the trialist on:
    1. How the RCT reported on the follow-up status for each category;
    2. How the RCT handled each category.

Based on these two criteria, authors of systematic reviews can judge whether participants belonging to each specific category have either (1) definitely missing data, (2) potentially missing data, or (3) definitely not missing data.

If trialists explicitly reported that these participants were followed-up, then systematic review authors should count them as definitely not having missing data. If trialists explicitly reported that participants were not followed-up, then systematic review authors should count them as definitely having missing data.

When trial authors do not explicitly report follow-up status - the case in 45% of RCTs (2)- systematic review authors should check how authors handled these categories. If such participants were excluded from the trial analysis (i.e., excluded from the denominator and numerator), then the reviewers should consider them definitely missing. Participants for whom the trialists imputed outcomes could be considered as having definitely missing data. However, the systematic reviewers should not treat them as missing data unless it is possible to obtain the number of observed/actual events (i.e., excluding imputed events) in order to avoid double counting. If it was unclear how primary study investigators handled participants with unclear follow-up status - the case in 52% of RCTs (2)- then it would be best to count them as potentially missing data.

Besides common situation mentioned in the table 1, we propose specific considerations for the following situations:

- Participants who are 'ineligible or mistakenly randomized' may be considered as appropriately excluded by the trialists if information about ineligibility was available at randomization and those making the decision regarding exclusion were blinded to allocation (30). Similarly, those who 'did not receive first dose/treatment', or who are

'ineligible due to early occurrence of outcome' may be considered as appropriately excluded by the trialists. Under these conditions, systematic review authors should not count such participants as having missing data. However, if either of these two conditions is not satisfied, the exclusion is considered inappropriate and these participants might have missing data, depending on how trialist report on their follow-up status.

- 'Experienced adverse events': when follow-up status of participants belonging to this category is not explicitly reported, we suggest that systematic review authors assume that they were followed-up and consequently do not have missing data.

Discussion

Summary

We present guidance for authors of systematic reviews on how to identify (and classify) participants with missing outcome data in reports of RCTs, and how to deal with presentations or descriptions that leave uncertainty as to the number of patients with missing data. Our approach uses categories of participants described in RCT reports, and who might have missing data, and relies on how trial authors report on those categories and handle them in their analyses (Table 2).

The guidance proposed in this paper complements existing guidance on handling and assessing risk of bias associated with missing data (6-14, 31-34); however, for their primary analysis, systematic reviewers must choose between two options: use either 'definitely missing data' or 'total possible missing data'. Review authors also need to choose a method for handling missing data in the meta-analysis (6). To test the robustness of the analysis that follows from these choices, the authors could explore sensitivity analyses using alternatives for identifying participants with missing outcome data and for handling missing data. Using the 'total possible missing data' (compared with using 'definitely missing data') in the primary analysis will yield a less precise pooled effect that is also less robust when subjected to sensitivity analyses. The main advantage of using the 'total possible missing data' (compared with using 'definitely missing data') is increased confidence in the results if the pooled effect estimate is found to be robust.

Strengths and limitations

To our knowledge, this is the first guidance for systematic review authors on how to identify participants with missing outcome data in reports of RCTs. The guidance is structured, transparent, and hopefully easy to implement. We labeled the categories to reflect and capture the wording used in RCT reports with which systematic review authors have to deal. We built the guidance based on extensive methodological work on the topic of missing data. We refined our recommendations using an iterative process during which we applied the categorization of participants who might have missing data to samples of RCTs and systematic reviews (a total of 1,073 RCTs and 302 systematic reviews).

One limitation is that our proposed approach to judging of data missingness did not benefit from as much validation as did the categorization of participants who might have missing data. However, our approach to judging of data missingness is consensus-based, and builds on the relevant data on the subject by Marciniak et al (21). Marciniak et al. compared the loss to follow-up rates in published reports of RCTs of oral antithrombotic, agents with loss to follow-up rates calculated based on more detailed documents made available to the Food and Drug Administration (FDA) for the same RCTs (21). They found a large discrepancy between the median of published rate of all ‘‘missing follow-up categories’’ (0.9%), and the median of the FDA-calculated loss to follow-up rates (13%). This suggests that missing data might be more frequent than what is explicitly reported and published.

The broader approach to addressing missing data

A broader approach to addressing missing data in trials includes (in addition to identifying and handling missing data) the avoidance and better reporting of missing data, as well as sharing individual participant data. Indeed, the best way to address missing data in RCTs is to minimize the extent of - if not avoid- missing data. Many strategies have been suggested to improve retention in RCTs (e.g., monetary incentives) (35-38). Similarly, trial methodologists have proposed strategies to improve the reporting of missing data (see box 3) (3, 22, 39, 40).

Box 3: Proposed approaches for reporting missing data

- Report methods used to prevent missing data;
- Report number of participants with missing data for each outcome, by study arm, and by time frame if relevant;
- Report rates of missing data by trial arms;
- Report a flow diagram of participants;
- Report any differences between baseline characteristics of participants with and without missing data;
- Report the reasons for missing data (refrain from reporting more than one category lumped together and 'other' category as they create further confusion and uncertainty);
- Report method(s) for handling missing data in analysis;
- Report results of any sensitivity analyses to assess the associated risk of bias;
- Discuss implication of missing data on interpreting the results.

If trial authors make individual participant data of their RCTs publicly available, systematic review authors will no longer need to make judgements on the extent of missing data of participants. Both the World Health Organization (WHO) and the International Committee of Medical Journal Editors (ICMJE) have emphasized the importance of sharing RCT data (41, 42). However, a recent survey of RCTs published in the BMJ and PLOS Medicine subsequent to the adoption of data sharing policies by these journals found that less than 50% of RCTs met criteria for data availability (41).

Short of eliminating missing data or sharing individual patient data, improved reporting of group level data can facilitate the handling of missing data in meta-analyses. Ideally, trialists would report, in a standardized data file compatible with meta-analysis software, the number of participants randomized, the number of participants with missing data, and the number of events for each outcome. We acknowledge that trialists are limited by word count in their journal publications, and that such information may appear in an appendix.

Implication for practice

As authors of systematic reviews will always face missing data in trials included in meta-analyses, we suggest the following stepwise approach to deal with missing data:

1. If trialists fail to provide the data in their trial report, request missing data by outcome and information on how they dealt with them in the analysis;
2. If the trialists do not provide sufficient information, follow the guidance suggested in this paper to identify participants who have either definitely or potentially missing data;
3. Assess risk of bias associated with missing data trial following GRADE guidance (8);
4. Report on all the above steps.

Implications for future research

Since this guidance has not been validated yet, it would be optimal to verify whether the categories judged to have missing actually have missing data, e.g., by comparing reported group-level data with individual participant data. Ultimately, the comparison of the approach against individual patient data would secure the validity of the approach.

Figures, tables, and supplementary data

Legends

Figure 1: Definitions of missing data

Table 1: Categories of participants described in RCT reports who might have missing data

Figure 2: How authors of systematic reviews could judge the missingness of outcome data in reports of randomized controlled trials

Table 2: Judging of outcome data missingness based on the reporting and handling in the trial of categories of participants that might have missing data.

Figure 1: Definitions of missing data

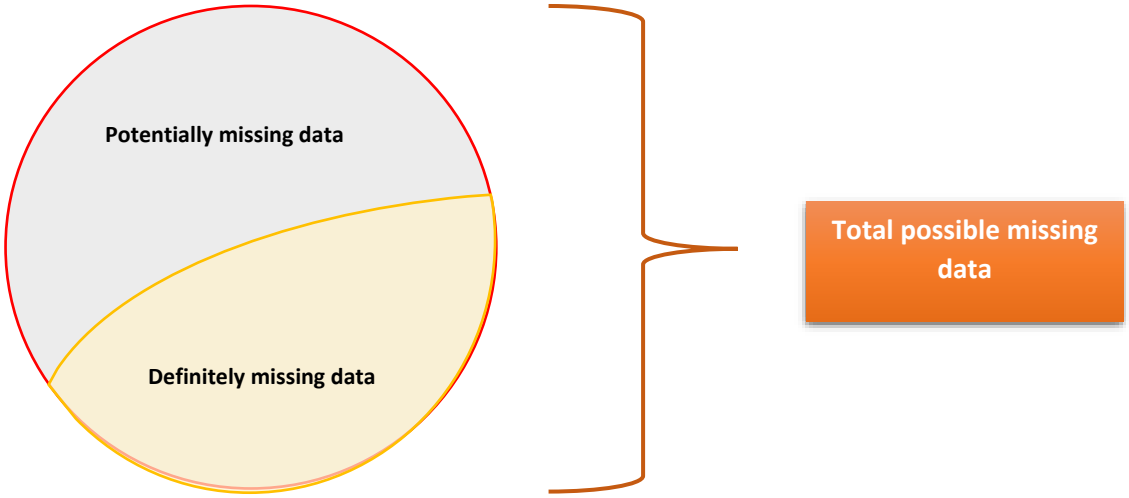
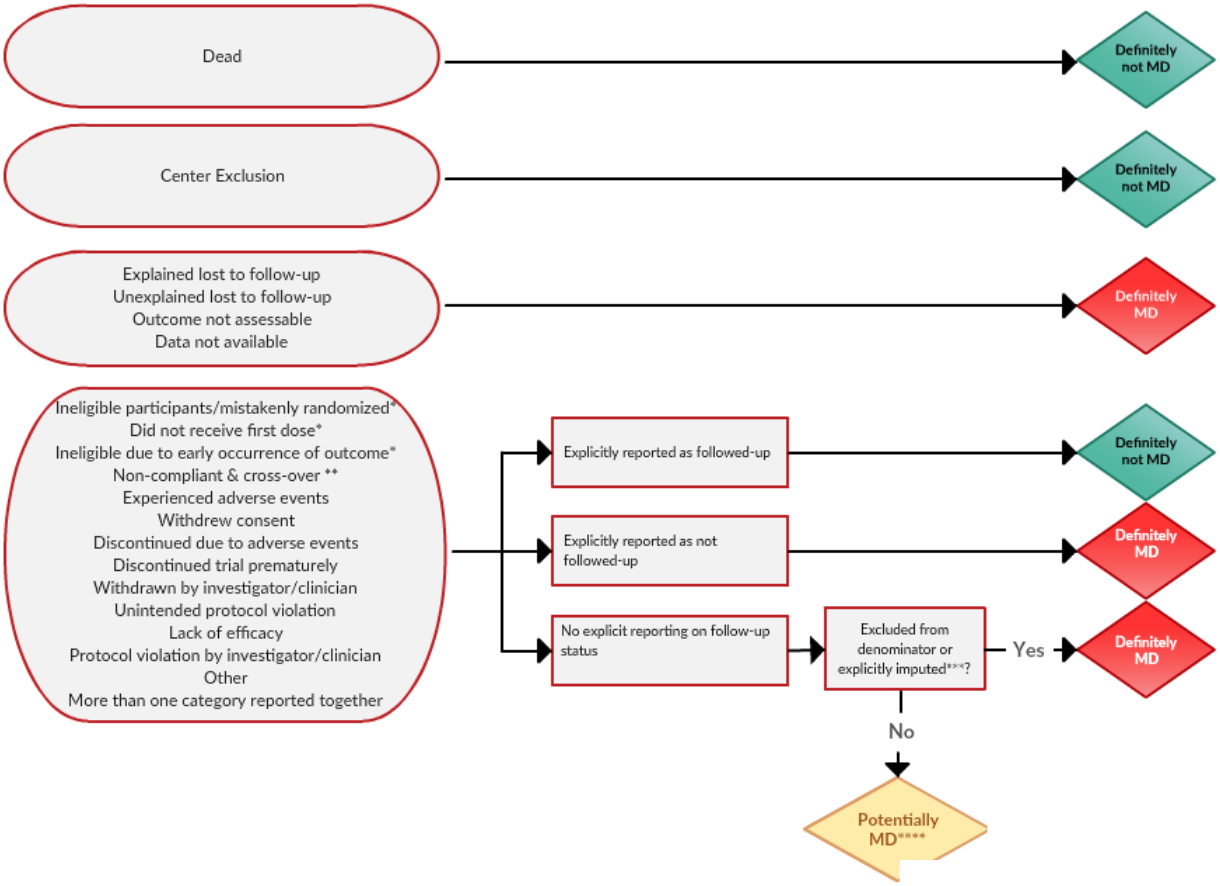


Table 1: Categories of participants described in RCT reports who might have missing data

Category of participants that might have missing data	Description of the category
Explained lost to follow-up	Participants described as lost to follow-up, and trialists provided an explanation, e.g., relocated to a different country
Unexplained lost to follow-up	Participants described as lost to follow-up, and trialists did not provide an explanation
Outcome not assessable	Data of a certain outcome for a number of participants is not available because the outcome adjudicators could not assess their outcome. For example, venography could not be done for a number of participants
Data not available	Participants who are still part of the RCT, however due to incomplete or missing record, the outcome data of these participant are missing
Ineligible or mistakenly randomized	Participants who, subsequent to randomization, are either found not to have the condition of interest (e.g. are not pregnant in an RCT among pregnant women), or did not undergo a procedure for which the intervention is intended (e.g. did not undergo surgery in an RCT of postoperative thromboprophylaxis)
Did not receive first dose/treatment	Participants who did not receive the 'first dose' of the intervention to which they were randomized
Ineligible due to early occurrence of outcome	Participants who were eligible at baseline then developed the outcome of interest soon after enrollment. These are considered ineligible if the trialists judge that the occurrence of the outcome cannot be related to the intervention of interest

Experienced adverse events	Participants who developed adverse events but without clear indication whether or not they discontinued the RCT
Non-compliant	Participants who were non-adherent or otherwise violated the protocol
Cross-over	Participants randomized to one arm, but who received the intervention meant for another treatment arm
Withdrew consent	Participants who withdraw their consent to participate in the RCT
Discontinued due to adverse events	Participants who discontinued the RCT due to adverse events
Discontinued trial prematurely	Participants who left the RCT but for whom a reason for discontinuation was not provided
Withdrawn by investigator/clinician	Participants who left the RCT through a decision made by the investigator or clinician (e.g., due to medical necessity)
Unintended protocol violation	Participants who left the RCT due a protocol violation for which they are not responsible (e.g., unavailability of hospital beds)
Lack of efficacy	Participants who left the RCT because they perceived no benefits from the intervention they were randomized to
Protocol violation by investigator/clinician	Investigator/clinician violated the protocol (e.g., change the intended intervention) due to a medical reason
More than one category reported together	The number refers to participants belonging to two or more of the above categories
Other	Reason different from the above

Figure 2: How authors of systematic reviews could judge the missingness of outcome data in reports of randomized controlled trials



Abbreviation: *FU: follow-up; MD: missing data.*
**Participants belonging to these categories not meeting the conditions for appropriate exclusion (see text).*
*** When both intention to treat analysis and per-protocol analysis are reported, assume that participants belonging to the categories are followed-up and consequently do not have missing data.*
****Participants for whom the trialists imputed outcomes could be considered as having definitely missing data. However, the systematic reviewers should not treat them as missing data unless it is possible to obtain the number of observed/actual events (i.e., excluding imputed events) in order to avoid double counting.*
*****When follow-up status is not explicitly reported in the RCT report, assume that they were followed-up and consequently do not have missing data.*

Table 2: Judging of outcome data missingness based on the reporting and handling in the trial of categories of participants that might have missing data.

Judging of outcome data missingness	Categories of participants that might have missing data
Definitely not missing data	• Participants explicitly reported as followed-up • Participants who died during the trial • Participants belonging to centers that were excluded

Definitely missing data	• Participants explicitly reported as not followed-up;
	• Participants with unclear follow-up status and:
	  ○ Excluded from the denominator of the analysis (i.e., complete case analysis), or
	  ○ Included in the denominator of the analysis and their outcomes were explicitly stated to be imputed. However, the systematic reviewers should not treat them as missing data unless it is possible to obtain the number of observed/actual events (i.e., excluding imputed events) in order to avoid double counting.

Potentially missing data	• Participants with unclear follow-up status (e.g., included in the denominator of the analysis and their outcomes were not explicitly stated to be imputed)
Total possibly missing data	• Participants who have either definite or potential missing data

References

1. Kahale LA DB, Brignardello-Petersen R, Agarwal A, Mustafa RA, Kwong J, Neumann I, Li L, Lopes LC, Briel M, Busse JW, Iorio A, Vandvik PO, Alexander PE, Guyatt G, Akl EA. Systematic reviews do not adequately report, or address missing outcome data in their analyses: a methodological survey. Journal of Clinical Epidemiology. 2018.
2. Kahale LA, Diab B, Khamis AM, Chang Y, Lopes LC, Agarwal A, et al. Potentially missing data was considerably more frequent than definitely missing data in randomized controlled trials: A methodological survey. Journal of Clinical Epidemiology. 2018.
3. Akl EA, Kahale LA, Ebrahim S, Alonso-Coello P, Schunemann HJ, Guyatt GH. Three challenges described for identifying participants with missing data in trials reports, and potential solutions suggested to systematic reviewers. J Clin Epidemiol. 2016.
4. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al. The prevention and treatment of missing data in clinical trials. N Engl J Med. 2012;367(14):1355-60.
5. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393.
6. Akl EA, Johnston BC, Alonso-Coello P, Neumann I, Ebrahim S, Briel M, et al. Addressing dichotomous data for participants excluded from trial analysis: a guide for systematic reviewers. PLoS One. 2013;8(2):e57132.
7. 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(9):1014-21 e1.
8. 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.
9. Higgins JP, White IR, Wood AM. Imputation methods for missing outcome data in meta-analysis of clinical trials. Clin Trials. 2008;5(3):225-39.
10. 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.
11. 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(12):2062-80.
12. 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(5):711-27.
13. White IR, Welton NJ, Wood AM, Ades AE, Higgins JP. Allowing for uncertainty due to missing data in meta-analysis--part 2: hierarchical models. Stat Med. 2008;27(5):728-45.
14. Higgins JPT SJ, Page MJ, and Sterne JAC. Chapter 8: Assessing risk of bias in a randomized trial. In: Draft version (16 September 2018) for inclusion in Higgins JPT TJ, Chandler J, Cumpston MS, Li T, Page MJ, Welch V, editor. Cochrane Handbook for Systematic Reviews of Interventions. London: The Cochrane Collaboration; 2018.
15. Hajek P, West R. Commentary on Smolkowski et al.(2010): Why is it important to assume that non-responders in tobacco cessation trials have relapsed? Addiction. 2010;105(6):1016-7.
16. West R, Hajek P, Stead L, Stapleton J. Outcome criteria in smoking cessation trials: proposal for a common standard. Addiction. 2005;100(3):299-303.

17. Foulds J, Stapleton J, Hayward M, Russell MA, Feyerabend C, Fleming T, et al. Transdermal nicotine patches with low-intensity support to aid smoking cessation in outpatients in a general hospital: a placebo-controlled trial. Archives of family medicine. 1993;2(4):417.
18. Jackson D, White IR, Mason D, Sutton S. A general method for handling missing binary outcome data in randomized controlled trials. Addiction. 2014;109(12):1986-93.
19. Spineli LM. Missing binary data extraction challenges from Cochrane reviews in mental health and Campbell reviews with implications for empirical research. Res Synth Methods. 2017;8(4):514-25.
20. Mavridis D, White IR. Dealing with missing outcome data in meta-analysis. Research synthesis methods. 2019.
21. Marciniak TA, Cherepanov V, Golukhova E, Kim MH, Serebruany V. Drug discontinuation and follow-up rates in oral antithrombotic trials. JAMA internal medicine. 2016;176(2):257-9.
22. Akl EA, Shawwa K, Kahale LA, Agoritsas T, Brignardello-Petersen R, Busse JW, et al. Reporting missing participant data in randomised trials: systematic survey of the methodological literature and a proposed guide. BMJ Open. 2015;5(12):e008431.
23. 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.
24. Zhang Y, Alyass A, Vanniyasingam T, Sadeghirad B, Flórez ID, Pichika SC, et al. A systematic survey of the methods literature on the reporting quality and optimal methods of handling participants with missing outcome data for continuous outcomes in randomized controlled trials. Journal of clinical epidemiology. 2017;88:67-80.
25. Zhang Y, Florez ID, Colunga Lozano LE, Aloweni FAB, Kennedy SA, Li A, et al. A systematic survey on reporting and methods for handling missing participant data for continuous outcomes in randomized controlled trials. J Clin Epidemiol. 2017;88:57-66.
26. 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.
27. 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.
28. Ebrahim S, Johnston BC, Akl EA, Mustafa RA, Sun X, Walter SD, et al. Addressing continuous data measured with different instruments for participants excluded from trial analysis: a guide for systematic reviewers. J Clin Epidemiol. 2014;67(5):560-70.
29. Noordzij M, Leffondré K, van Stralen KJ, Zoccali C, Dekker FW, Jager KJ. When do we need competing risks methods for survival analysis in nephrology? Nephrology Dialysis Transplantation. 2013;28(11):2670-7.
30. Fergusson D, Aaron SD, Guyatt G, Hébert P. Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis. BMJ: British Medical Journal. 2002;325(7365):652.
31. 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(5):721-41.

32. Spineli LM, Higgins JP, Cipriani A, Leucht S, Salanti G. Evaluating the impact of imputations for missing participant outcome data in a network meta-analysis. Clinical Trials. 2013;10(3):378-88.
33. Altman DG. Missing outcomes in randomized trials: addressing the dilemma. Open Med. 2009;3(2):e51-3.
34. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials–a practical guide with flowcharts. BMC medical research methodology. 2017;17(1):162.
35. Brueton VC, Tierney J, Stenning S, Harding S, Meredith S, Nazareth I, et al. Strategies to improve retention in randomised trials. The Cochrane database of systematic reviews. 2013(12):1.
36. Dziura JD, Post LA, Zhao Q, Fu Z, Peduzzi P. Strategies for dealing with missing data in clinical trials: from design to analysis. Yale J Biol Med. 2013;86(3):343-58.
37. Hollestein L, Carpenter J. Missing data in clinical research: an integrated approach. British Journal of Dermatology. 2017;177(6):1463-5.
38. Little R, Cohen M, Dickersin K, Emerson S, Farrar J, Neaton J, et al. The design and conduct of clinical trials to limit missing data. Statistics in medicine. 2012;31(28):3433-43.
39. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux P, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Bmj. 2010;340:c869.
40. National Research Council. The Prevention and Treatment of Missing Data in Clinical Trials. Washington (DC): The National Academic Press; 2010.
41. Naudet F, Sakarovitch C, Janiaud P, Cristea I, Fanelli D, Moher D, et al. Data sharing and reanalysis of randomized controlled trials in leading biomedical journals with a full data sharing policy: survey of studies published in The BMJ and PLOS Medicine. bmj. 2018;360:k400.
42. Taichman DB, Sahni P, Pinborg A, Peiperl L, Laine C, James A, et al. Data sharing statements for clinical trials: a requirement of the International Committee of Medical Journal Editors. Annals of internal medicine. 2017;167(1):63-5.

Chapter 5

Potential impact of missing outcome data on treatment effects in systematic reviews: an imputation study

Kahale LA, Khamis AM, Diab B, Chang Y, Lopes LC, Agarwal A, Li L, Mustafa R, Koujanian S, Waziry R, Busse JW, Dakik A, Schünemann H, Hooft L, Guyatt G, Scholten RJPM, and Akl EA BMJ (submitted)

Abstract

Background:

Missing data for the outcomes of participants in randomized controlled trials (RCTs) may introduce bias in systematic reviews. To assess that risk of bias, the GRADE working group recommends challenging the robustness of the meta-analysis effect estimate by conducting sensitivity analyses with different methods of handling missing data.

Objective:

To assess risk of bias associated with missing outcome data in systematic reviews we calculated (1) the percentage of meta-analyses that lost statistical significance with each of these methods; (2) the percentage of meta-analyses that changed direction of effect, and (3) the median change of effect estimates across meta-analyses.

Methods:

We selected systematic reviews that included a group-level meta-analysis with a statistically significant effect on a patient-important, dichotomous efficacy outcome. We conducted sensitivity analyses based on different methods of handling missing data. These included four commonly discussed but implausible assumptions (e.g., worst-case scenario) and four plausible assumptions for missing data based on the informative missingness odds ratio (IMOR) approach. For each method, we specifically calculated: (1) the percentage of meta-analyses that lost statistical significance, (2) the percentage of meta-analyses that qualitatively changed direction of effect, and (3) the median percentage change in the relative effect estimate when applying each assumption.

Results:

We included 100 systematic reviews with 653 RCTs. When applying the implausible but commonly discussed assumptions, 1% (best-case scenario) to 60% (worst-case scenario) of meta-analyses lost statistical significance, while 26% changed direction with the worst-case scenario. The median change in the relative effect estimate varied from 0% to 30.4%. When applying the plausible assumptions, 6% (least stringent IMOR) to 22% (most stringent IMOR) of meta-analyses lost statistical significance, while 2% changed

direction with the most stringent IMOR. The median percentage change in relative effect estimate varied from 1.4% to 7.0%.

Conclusion:

Even when applying plausible assumptions to the outcomes of participants with definite missing data, almost a quarter (22%) of meta-analyses lost statistical significance. Systematic review authors should present the potential impact of missing outcome data on their effect estimates and use this to inform their overall GRADE ratings of risk of bias and their interpretation of the results.

Introduction

Despite efforts to reduce their incidence [1], randomized controlled trials (RCTs) commonly suffer from missing outcome data. The percentage of RCTs with missing outcome data across six methodological surveys ranged from 63% to 100%, [2-7] and the average proportion of participants with missing data among trials reporting missing data ranged from 6% to 24% [2-8]. Among 235 RCTs with statistically significant results published in leading medical journals, one in three lost statistical significance when making plausible assumptions about the outcomes of participants with missing data [2]. Another study comparing different approaches to modeling binary outcomes with missing data in an alcohol clinical trial yielded different results with various amount of bias depending on the approach and missing data scenario [9].

The extent of missing outcome data in RCTs contributes to the risk of bias of meta-analyses including those RCTs. To explore the impact on risk of bias, the GRADE working group recommends conducting sensitivity analyses using assumptions regarding the outcomes of patients with missing outcome data [10]. No methodological study has thus far assessed the impact of different methods regarding missing data on the robustness of the pooled relative effect in a representative sample of systematic reviews.

One challenge when handling missing data is the lack of clarity in trial reports on whether participants have missing outcome data [13]. We recently published guidance for authors of systematic reviews on how to identify and classify participants with missing outcome data in reports of RCTs [14]. The Cochrane Handbook acknowledges that attempts to address missing data in systematic reviews are often hampered by incomplete reporting of missing outcome data by trialists [11]. A recent methodological survey among 638 RCTs reported that the median percentage of participants with unclear follow-up status was 9.7% (IQR 4.1%-19.9%) [8], and that when RCT authors explicitly reported not following-up participants, almost half did not specify how they handled missing data in their analysis.

Objectives

The objective of this study was to assess risk of bias associated with missing outcome data in systematic reviews by: (1) examining how different methods of handling missing data alter statistical significance of pooled effect estimates; (2) qualitatively changed the direction of effect; and (3) quantifying the change in effect estimate when applying different methods of handling missing outcome data.

Methods

Design overview

This study is part of a larger project examining methodological issues related to missing data in systematic reviews and RCTs [12]. Our published protocol includes detailed information on the definitions, eligibility criteria, search strategy, selection process, data abstraction and data analysis [12]. In the appendix (section 1), we present the deviations from the protocol and the corresponding rationale. This study did not involve human subjects and thus did not require ethical approval.

We defined missing data as outcome data for trial participants that are not available to authors of systematic reviews (from the published RCT reports or personal contact with RCT authors).

In the current study, we collected a random sample of 50 Cochrane and 50 non-Cochrane systematic reviews published in 2012 that reported a group-level meta-analysis of a patient-important dichotomous efficacy outcome, with a statistically significant effect estimate (the 'meta-analysis of interest') [13]. We used the term 'original pooled relative effect' to refer to the result of the meta-analysis as reported by the systematic review authors. For the individual RCTs included in the meta-analyses of interest [8], we abstracted detailed information relevant to the statistical analysis and missing data and conducted sensitivity meta-analyses based on nine different methods of handling missing data. Our outcomes were (1) the percentage of meta-analyses that lost statistical significance with each of these methods; (2) the percentage of meta-analyses that

changed direction of effect, and (3) the median change of effect estimates across meta-analyses.

Identifying which participants have missing data

As noted above, and since publication of our protocol, we published a guidance for authors of systematic reviews on how to identify and classify participants with missing outcome data in reports of RCTs depending on how trial authors report on those categories and handle them in their analyses (Table 1) [14]. The guidance includes a taxonomy of categories of RCT participants who might have missing data, along with a description of those categories (appendix section 2). The categorization reflects the wording used in RCT reports, i.e., the presentation that systematic review authors actually face. We used this guidance to judge the outcome data missingness of categories of participants that might have missing data (i.e., classify whether they have definite missing, potential missing data, or no missing data).

Data abstraction

A group of 11 reviewers trained in health research methodology conducted data abstraction, independently and in duplicate. Reviewers met regularly with a core team (EAA, LAK, BD, and AD) to discuss progress and challenges and develop solutions. We used a pilot tested standardized data abstraction form hosted in an electronic data capture tool, REDCap [15]. All reviewers underwent calibration exercises prior to data abstraction to promote reliability, and a senior investigator served as a third independent reviewer for resolving disagreements.

We abstracted the following data of each eligible meta-analysis: (1) original (i.e., published) pooled relative effect, i.e., pooled relative effect measure (RR or OR) and the associated 95% confidence interval (CI); (2) the analysis model used (i.e., random effects or fixed effect); and (3) statistical method used for pooling (e.g., Mantel-Haenszel or Peto).

For each RCT, we abstracted the following data for each study arm: (1) number of participants randomized; (2) number of events; (3) number of participants that had

definite missing data (according to the suggested guidance on identifying participants with missing data [14]); and (4) number of participants that have potential missing data.

Data Analysis

We conducted a descriptive analysis of study characteristics of eligible systematic reviews and their associated RCTs with SPSS statistical software, version 21.0 [16]. For categorical variables, we reported frequencies and percentages. For continuous variables that were not normally distributed we used median and interquartile range (IQR).

In order to explore the robustness of pooled effect estimates reported by systematic reviews, we conducted a number of sensitivity analyses based on nine different methods of handling missing data using Stata software release 12 [17] (see table 2):

- Complete case analysis (CCA)[10, 18];
- Four implausible but commonly discussed assumptions: best-case scenario, none of participants with missing data had the outcome, all participants with missing data had the outcome, worst-case scenario;
- Four plausible assumptions using increasingly stringent values of informative missing odds ratio (IMOR) in the intervention arm [19, 20]. IMOR describes the ratio of odds of the outcome among participants with missing data to the odds of the outcome among observed participants. In other words, to obtain the odds among participants with missing data, one multiplies the odds among the observed participants with a stringent value (e.g., 1, 2, 3, 5).

Imputing events consists of including participants with missing data in the denominator and making assumptions about their outcomes in the numerator. This approach may lead to imputing a number of events 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 [19, 20, 22]. As the purpose of this project was to help in judging the risk of bias associated with missing data in systematic reviews rather than to generate alternative best estimates of intervention effect, we did not consider the uncertainty when conducting the sensitivity analyses.

The following is a detailed step-by-step description of the analytical approach executed in one command (metamiss [21]) by Stata release 12[17] for each sensitivity analysis.

First, we re-calculated each meta-analysis of interest, for all 100 systematic reviews, using each method to address missing outcome data to generate different sensitivity analysis pooled effects along with their 95% CIs. We used the same relative effect measure (RR or OR), the same analysis model (random effects or fixed effect), and the same statistical method (e.g., Mantel-Haenszel) as the original meta-analysis of interest.

Second, across all included meta-analyses and for each method, we explored the impact of the revised meta-analysis in regard to the following outcomes:

1. The percentage of meta-analyses for which the 'sensitivity analysis pooled relative effect (assumption)' lost statistical significance compared to the 'sensitivity analysis pooled relative effect (CCA)'. For this analysis, we restricted the sample to the meta-analyses that remained significant under the CCA method.
2. The percentage of meta-analyses for which the 'sensitivity analysis pooled relative effect (assumption)' changed direction compared to the 'sensitivity analysis pooled relative effect (CCA)'. The direction could change either from favoring the intervention to favoring the control or vice versa.
3. Change in the relative effect estimate: To quantify the percentage change in relative effect estimate between the 'sensitivity analysis pooled effect estimate (assumption)' and the 'sensitivity analysis pooled effect estimate (CCA)', we applied the following formula (see further statistical notes in the appendix section 3):

$$\frac{\text{Sensitivity analysis pooled relative effect (assumption)} - \text{Sensitivity analysis pooled relative effect (CCA)}}{\text{Sensitivity analysis pooled relative effect (CCA)}}$$

We calculated specifically:

- The percentage of meta-analyses with change of relative effect estimate (by direction) between the 'sensitivity analysis pooled effect estimate (assumption)' and the 'sensitivity analysis pooled effect estimate (CCA)';

- The median and interquartile range (IQR) for the change in relative effect estimate (stratified by direction of change).

This relative change could be an increase or a reduction in effect. For example, a relative increase in relative risk of 25% for the 'sensitivity analysis pooled effect estimate (worst case scenario)' over the 'sensitivity analysis pooled effect estimate (CCA)' implies that:

- for a relative risk of 0.8 with CCA, the relative risk for the worst-case scenario would be 1;
- for a relative risk of 1.6 with CCA, the relative risk for the worst-case scenario would be 1.2.

We reproduced the analyses for outcomes '1' and '2' (i.e., loss of statistical significance and changing direction) comparing the 'sensitivity analysis pooled relative effect' to the 'original pooled relative effect'. In addition, we conducted the above analyses twice: first considering participants with definite missing outcome data, second considering participants with total possible missing outcome data (see Table 1).

Results

Study characteristics of included meta-analyses

We previously reported on the details of the 100 eligible systematic reviews, [13] and the 653 RCTs they considered [8]. Table 3 summarizes the characteristics of the 100 included systematic reviews and their corresponding meta-analyses. The majority reported on a pharmacological outcome (61%) and a non-active control (55%), assessed a morbidity outcome (56%), reported an unfavorable outcome (73%), used the risk ratio (61%), and applied a fixed-effect analysis model (57%) and Mantel- Haenszel statistical methods (77%). The median number of RCTs per meta-analysis was 6 with an IQR of 3-8. Eight meta-analyses included RCTs that reported no missing data.

Missing data in the RCTs

Four hundred of the 653 RCTs (63%) mentioned in their results at least one of the pre-defined categories of participants who might have missing outcome data. Among those

400 RCTs, the median percentage of participants with definite missing outcome data was 5.8% (IQR 2.2%-14.8%), with potential missing outcome data 9.7% (IQR 4.1%-19.9%), and with total possible missing data 11.7% (IQR 5.6%-23.7%). Only three RCTs described a mechanism of missingness (e.g., missing at random).

Loss of statistical significance and change of direction

As previously noted, 87 out the 100 meta-analyses maintained statistical significance under the CCA method. Figure 1 shows the number of meta-analyses that lost significance when comparing the 'sensitivity analysis pooled relative effect (assumption)' to the 'sensitivity analysis pooled relative effect (CCA)' for each assumption and considering participants with definite missing data. For the four implausible but commonly discussed assumptions, the percentage of meta-analyses that lost significance ranged from 1% (best-case scenario and none of the participants with missing data had the event) to 18% (all participants with missing data had the event) to 60% (worst-case scenario). For the plausible assumptions based on IMOR, the percentage of meta-analyses that lost significance ranged from 6% (least stringent assumption IMOR 1.5) to 22% (most stringent assumption IMOR 5).

The percentage of meta-analyses that changed direction with the two extreme assumptions was 26% for the worst-case scenario and 2% for the most stringent assumption IMOR 5.

We present in the appendix (sections 4 and 5 respectively) the results of meta-analyses that (1) lost significance and (2) changed direction:

- when considering participants with total possible missing data;
- when comparing the 'sensitivity analysis pooled relative effect (assumption)' to the 'original pooled relative effect'.

Change in the relative effect estimate

Figure 2 shows the change in the relative effect estimate between the 'sensitivity analysis pooled effect estimate (assumption)' and the 'sensitivity analysis pooled effect estimate (CCA)', when considering participants with definite missing data.

For the four implausible but commonly discussed assumptions, the percentage of meta-analyses with increased relative effect estimate (shifted away from the null value of 1)

was 91% for the 'best case scenario' assumption, 25% for 'none of the participants with missing data had the event' assumption, and 17% for 'all participants with missing data had the event' assumption. The median increase in the relative effect estimate ranged from 0% for the 'worst case scenario' assumption to 18.9% (IQR 6.8%-38.9%) for the 'best case scenario' assumption. The percentage of meta-analyses with reduced relative effect estimate (shifted closer towards the null value of 1) was 90% for the 'worst case scenario' assumption, 38% for 'none of the participants with missing data had the event' assumption, and 75% for 'all participants with missing data had the event' assumption. The median reduction in the relative effect estimate ranged from 0% for the 'best case scenario' assumption to 30.4% (IQR 10.5%-77.5%) for the 'worst case scenario' assumption.

For the plausible assumptions based on the IMOR, the percentage of meta-analyses with increased relative effect estimate was 85% for the least stringent assumption (IMOR 1.5) and 88% for the most stringent assumption (IMOR 5). The median reduction in relative effect estimate ranged from 1.4% (IQR 0.6%-3.9%) for the IMOR 1.5 assumption to 7.0% (IQR %2.7-18.2%) for the IMOR 5 assumption. We present in the appendix (sections 6 and 7 respectively) the details of the percentage change in the relative effect estimate: (1) when considering participants with total possible missing data; and (2) stratified by whether the estimate is less than or greater than 1 under the CCA, using either definite missing data or total possible missing data.

Discussion

Summary of findings

In the current study, we examined how different methods of handling missing data alter the statistical significance of pooled effect estimates of dichotomous outcomes. Even when applying plausible assumptions to the outcomes of participants with definite missing data, almost a quarter (22%) of meta-analyses lost statistical significance. When applying implausible but commonly discussed assumptions, the percentage of systematic reviews that lost significance was as high as 60% with the worst-case scenario.

We also quantified the change in the effect estimate when applying assumptions to the outcomes of participants with definite missing data. When applying plausible

assumptions to the outcomes of participants with definite missing data, the median change in relative effect estimate was as high as 7.0% (IQR 2.7%-18.2%). When applying implausible but commonly discussed assumptions, the median change in the relative effect estimate was as large as 30.4% (IQR 10.5%- 77.5%).

Strengths and limitations

This is the first study to assess the effect of using different assumptions (both the commonly discussed and more plausible) on a large number of published meta-analyses of patient-important outcomes addressing a wide range of clinical topics. Strengths of our study include a very detailed approach to assessing participants with missing data, and accounting for participants with 'potential missing data' in our analyses. We used two statistical approaches to assess the risk of bias associated with missing outcomes: loss of statistical significance and change in effect estimate. Although the former approach has been criticized as the basis for decision making, [23] we used it to assess the robustness of the meta-analysis effect estimate (i.e., when conducting sensitivity meta-analyses using different methods of handling missing data). We are confident that if results lose statistical significance, the certainty in evidence should be rated down due to risk of bias. The 'change in effect estimate' approach has its own limitation in terms of interpretation (cutoff for topic-specific minimally important difference that would vary across a wide range of topics and outcomes).

A limitation of our study is that we considered only dichotomous outcome data; methods for handling missing continuous data are different and our findings may not be generalizable to systematic reviews of continuous outcomes [24, 25]. Our sample consisted of systematic reviews that were published in 2012 and these may not reflect more current reviews; however, recent surveys have found the reporting, handling, and assessment of risk of bias in relation to missing data has not improved over this period of time [7, 26-28]. Another limitation is our focus on systematic reviews with statistically significant results. While this prevented us from assessing the change in effect estimate for meta-analyses with non-statistically significant results, it allowed us to focus on reviews that are more likely to influence clinical practice.

Interpretation of findings

Almost a quarter of meta-analyses lost significance when using a conservative approach to test their robustness (i.e., applying plausible assumptions to the outcomes of participants with definite missing data). When using the same conservative approach, up to a quarter of meta-analyses had a change of at least 18% in their relative effect estimates (based on the 75th for IMOR 5- refer to median and IQR of IMOR 5 in figure 1). These findings mean that a substantive percentage of meta-analyses is at serious risk of bias associated with missing outcome data. Findings such as these should lead systematic review authors to rate down the certainty of evidence for risk of bias. Our results highlight the importance of minimizing missing data for clinical trials, [29, 30] better reporting and handling of missing data [8, 13, 31].

The assumptions 'all participants with missing data had the event' and 'none of the participants with missing data had the event', may either increase or decrease the effect estimate. Thus, these two methods do not allow assessing the robustness of the effect estimate. By design, the 'best case scenario' assumption shifts the effect estimate in the opposite direction of challenging the robustness and should not be used for that purpose. The worst-case scenario consistently challenges the robustness of the effect estimate by shifting it towards the null effect, but the implausibility of its underlying assumptions makes it a poor choice for sensitivity analyses. Only if the effect estimate is robust to the worst-case scenario one can conclude that the evidence is at low risk of bias due to missing data.

There is a growing experience and acceptance for using plausible assumptions as they have face validity [32, 33]. The advantage of the IMOR approach is that it provides a tool where the review authors can challenge the robustness of effect estimates by applying increasingly stringent assumptions (i.e., 2, 5). This approach allows the review authors to choose the IMOR value based on their clinical judgment. We did not use IMOR 1 because it provides the same effect estimate as the CCA while narrowing the confidence interval.

Conclusion

The findings of this study show the potential impact of missing data on the results of systematic reviews. This has implications on both assessing the risk of bias associated

with missing outcome data, and the need to reduce the extent of missing outcome data in clinical trials.

Systematic review authors should present the potential impact of missing outcome data on their effect estimates and, when these suggest lack of robustness of the results, rate down the certainty of evidence for risk of bias. For practical purposes, authors of systematic reviews might wish to use statistical software that allow running assumptions about missing data (e.g., Stata). As for users of the medical literature, there is a need to come up with a rule of thumb on how to judge risk of bias associated with missing outcome data at the trial level. Such rule of thumb would account for factors such as:

- Percentage of missing data per study arm;
- Ratio of missing data to event rate per arm (i.e., the higher the ratio, the larger the change);
- Fragility of statistical significance (i.e., borderline significance);
- Magnitude of the effect estimate (i.e. the larger the effect estimate, the smaller the change);
- Duration of follow-up (i.e., the longer the duration of follow-up, the higher the percentage of missing data).

We acknowledge that assessing the impact of missing data with loss of statistical significance might be insensitive. Thus, when using this approach and statistical significance is lost, rate down the certainty for risk of bias associated with missing data. If statistical significance is not lost, it might be valuable to then evaluate the change in effect estimate to assess whether the relative effect goes from an important to an unimportant effect. If the latter happens, then rate down the certainty for risk of bias associated with missing data. However, the judgment of whether the change in effect estimate is clinically significant requires using minimal clinically important difference, which varies by clinical question. Thus, it would be ideal to reproduce this study in specific field(s) of medical science with clearly defined minimal clinically important differences.

Future research could also validate some of the findings of this study. For example, one could reproduce this study using individual participant data meta-analyses and compare

its findings to the current study. Also, individual participant data meta-analyses would allow testing other imputation methods (e.g., multiple imputations).

Figures, tables, and supplementary data

Legends

Table 1: Judging of outcome data missingness based on the reporting and handling of categories of participants that might have missing data

Table 2: List and description of the different methods of handling missing data

Table 3: General characteristics of the included systematic reviews and their meta-analyses (N=100)

Figure 1: Results of meta-analyses that lost significant when considering participants with definite missing data and comparing the 'sensitivity analysis pooled relative effect (assumption)' to the 'sensitivity analysis pooled relative effect (CCA)' (n=87 systematic reviews that maintained statistical significance under the CCA)

Figure 2: Change of relative effect estimate (by direction) between the 'sensitivity analysis pooled effect estimate (assumption) and the 'sensitivity analysis pooled effect estimate (CCA)' when considering participants with definite missing data. Bars in the upper part of the figure represent the percentage of meta-analyses with change of relative effect estimate (by direction). The numerical values in the bottom part represent the median (IQR) for, respectively, the increase and decrease in relative effect estimate (N=100)

Supplementary table 1: Deviations from the protocol and the corresponding justification for each deviation

Supplementary table 2: Categories of RCT participants who might have missing data

Supplementary statistical notes: Data analysis

Supplementary results 1: The percentage of meta-analyses for which the 'sensitivity analysis pooled relative effect' (1) lost statistical significance and (2) changed direction compared to the 'sensitivity analysis pooled relative effect (CCA)' when considering total possible missing data.

Supplementary results 2: The percentage of meta-analyses for which the 'sensitivity analysis pooled relative effect' (1) lost statistical significance and (2) changed direction compared to the 'original pooled relative effect' when considering definite and total possible missing data.

Supplementary results 3: The percentage change in the relative effect estimate between the 'sensitivity analysis pooled effect estimate (assumption)' and the 'sensitivity analysis pooled effect estimate (CCA)', when considering participants with total possible missing data.

Supplementary table 3: Details of the percentage change in the relative effect estimate, stratified by whether the estimate is less than or greater than 1 under the complete case analysis (CCA), using either definite missing data or total possible missing data

Table 1: Judging of outcome data missingness based on the reporting and handling of categories of participants that might have missing data

Categories of participants	Judgment of outcome data missingness
• Participants explicitly reported as followed-up • Participants who died during the trial • Participants belonging to centers that were excluded	**Definitely not missing data**

• Participants explicitly reported as not followed-up;
• Participants with unclear follow-up status and:
  ○ excluded from the denominator of the analysis (i.e., complete case analysis), or
  ○ included in the denominator of the analysis and their outcomes were explicitly stated to be imputed
	Definite missing data

Participants with unclear follow-up status (e.g., included in the denominator of the analysis and their outcomes were not explicitly stated to be imputed)	**Potential missing data**
Participants who have either definite or potential missing data	**Total possible missing data**

Table 2: List and description of the different methods of handling missing data

Method of handling missing data	Handling participants with missing data in the numerator and denominator			
	Intervention arm		Control arm	
	Numerator	Denominator	Numerator	Denominator
Complete case analysis	Excluded	Excluded	Excluded	Excluded
Implausible but commonly used assumptions				
Best-case scenario [a]	Assumed that all had a favorable outcome	Included	Assumed that all had an unfavorable outcome	Included
None of the participants with missing data had the outcome	Assumed that all did not have the outcome	Included	Assumed that all did not have the outcome	Included
All participants with missing data had the outcome	Assumed that all had the outcome	Included	Assumed that all had the outcome	Included
Worst-case scenario [b]	Assumed that all had an unfavorable outcome	Included	Assumed that all had a favorable outcome	Included
Plausible assumptions [c]				
IMOR 1.5	IMOR 1.5 [d]	Included	IMOR 1	Included
IMOR 2	IMOR 2 [d]	Included	IMOR 1	Included
IMOR 3	IMOR 3 [d]	Included	IMOR 1	Included
IMOR 5	IMOR 5 [d]	Included	IMOR 1	Included

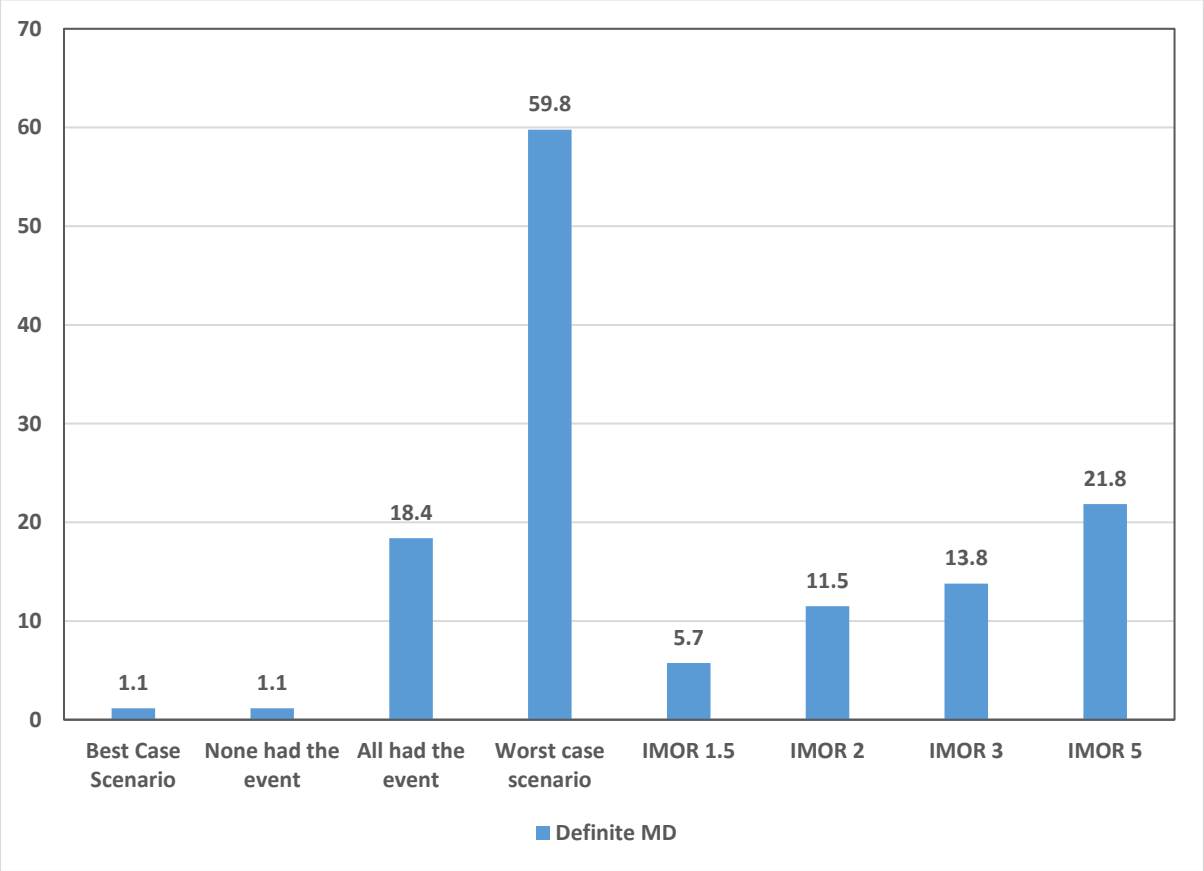
Abbreviations: *IMOR: informative missing odds ratio.*
[a] *When applying best-case scenario, we ensured it challenges the relative effect by shifting it away from the null value of no effect (see further statistical notes in the appendix section 3).*
[b] *When applying worst-case scenario, we ensured it challenges the relative effect by shifting it closer to the null value of no effect (see further statistical notes in the appendix section 3).*
[c] *We used the 'metamiss' command [20] to implement the IMOR assumptions in Stata.*
[d] *These calculations are applied when the relative effect is less than 1. When relative effect is greater than 1, the values for the IMOR are flipped between the intervention and control arm whereby it is 1 for the intervention arm. For example, when original relative effect is greater than 1, IMOR value for the intervention arm would be 1 and that of the control arm would be 5.*

Table 3: General characteristics of the included systematic reviews and their meta-analyses (N=100)

Characteristic	n (%)
Number of RCTs per meta-analysis (median (IQR))	6 (3 – 8)
Type of intervention	
Pharmacological	61 (61.0)
Surgery/invasive procedure	24 (24.0)
Other	15 (15.0)
Type of control	
Active: pharmacological	21 (21.0)
Active: surgery/invasive procedure	18 (18.0)
Non-active: no intervention/standard of care/placebo/sham	55 (55.0)
Other	6 (6.0)
Outcome category	
Mortality	21 (21.0)
Morbidity	56 (56.0)
Patient reported outcomes	23 (23.0)
Nature of outcome *	
Favorable	27 (27.0)
Unfavorable	73 (73.0)
Duration of outcome follow-up in months (mean, SD)	12.5 (23.1)
Effect measures reported	
Risk ratio (RR)	61 (61.0)
Odds ratio (OR)	39 (39.0)
Analysis model	
Random effects model (RE)	43 (43.0)
Fixed effect model (FE)	57 (57.0)
Statistical methods	
Mantel-Haenszel (MH)	77 (77.0)
Inverse variance (I-V)	4 (4.0)
Peto	7 (7.0)
Other	7 (7.0)
Not reported	5 (5.0)
Reported a handling method	
Complete case analysis	2 (2.0)
Assuming none of participants with missing data had the event	3 (3.0)
Assuming all of the participants with missing data had the event	2 (2.0)
Not reported	93 (93.0)

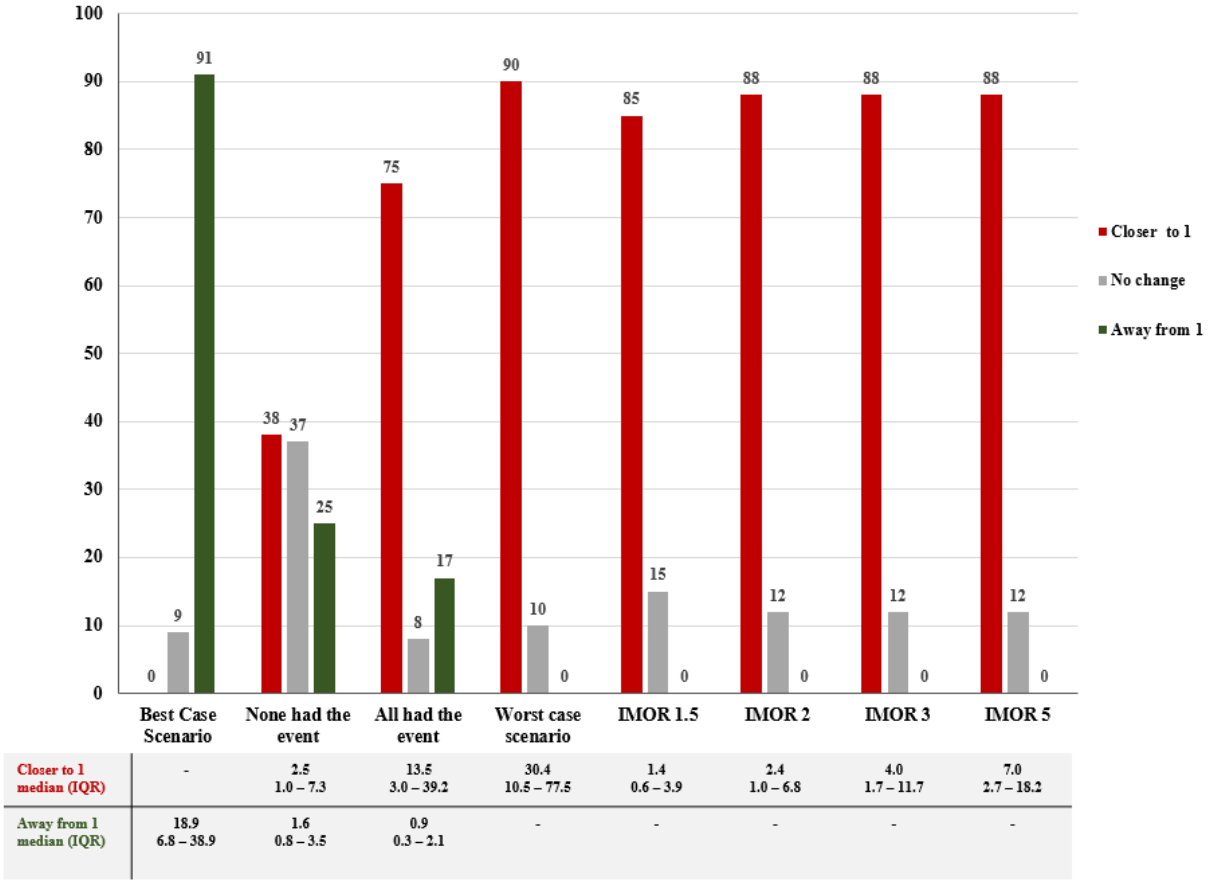
***Abbreviations**: FE: fixed effect; IQR: interquartile range; I-V: inverse variance; MH: Mantel-Haenszel; OR: odds ratio; RE: random effect; RCTs: randomized controlled trials; RR: risk ratio*
**Nature of the outcome refers to whether the outcome is negative (e.g., mortality) or positive (e.g., survival).*

Figure 1: Results of meta-analyses that lost significant when considering participants with definite missing data and comparing the 'sensitivity analysis pooled relative effect (assumption)' to the 'sensitivity analysis pooled relative effect (CCA)' (n=87 systematic reviews that maintained statistical significance under the CCA)



***Abbreviations**: IMOR: informative missing odds ratio; MD: missing data*

Figure 2: Change of relative effect estimate (by direction) between the 'sensitivity analysis pooled effect estimate (assumption) and the 'sensitivity analysis pooled effect estimate (CCA)' when considering participants with definite missing data. Bars in the upper part of the figure represent the percentage of meta-analyses with change of relative effect estimate (by direction). The numerical values in the bottom part represent the median (IQR) for, respectively, the increase and decrease in relative effect estimate (N=100)



Abbreviations: *IMOR: informative missing odds ratio; IQR: interquartile range; MD: missing data*

Supplementary table 1: Deviations from the protocol and the corresponding justification for each deviation

Item	What was stated in protocol	The deviation	Justification for the change
Selecting and reproducing the original meta-analysis of interest	'For each eligible meta-analysis, we will first attempt to reproduce the original analysis. When this analysis generates a different effect estimate that is not statistically significant, we will exclude the corresponding meta-analysis from this part of the study.'	We included the eligible meta-analyses without reproducing the original analysis.	When we attempted to reproduce the original analysis, we found it very challenging to figure out what data the systematic reviewers used in their analysis.
Assumed effect among participants with missing data relative to effect observed among followed-up participants	'We define $RI_{NotFU/FU}$ as the relative event incidence among those not followed-up relative to the event incidence among those followed-up'	Instead of using the $RI_{NotFU/FU}$, we used the informative missing odds ratio (IMOR) method which describes the relationship between the unknown odds of the outcome among participants with missing data and the known odds among observed participants.	We decided to use the IMOR because of it easily applied in Stata (metamiss command). On the other hand, $RI_{NotFU/FU}$ has not such command available. In addition, the two methods rendered the comparable results when applied on a sample of 52 meta-analyses.
Uncertainty associated with the imputed values	'We will apply to each of these assumptions and statistical	We did not take uncertainty into account while	As the purpose of this project was to help in judging the risk of bias

	approaches to take uncertainty into account.'	applying assumptions.	associated with missing data in systematic reviews rather than to generate alternative best estimates of intervention effect, we judged that the sensitivity analyses do not require taking uncertainty into account.

Supplementary table 2: Categories of RCT participants who might have missing data (71)

Category of participants that might have missing data	Description of the category
Explained lost to follow-up	Participants described as lost to follow-up, and trialists provided an explanation, e.g., relocated to a different country
Unexplained lost to follow-up	Participants described as lost to follow-up, and trialists did not provide an explanation
Outcome not assessable	Data of a certain outcome for a number of participants is not available because the outcome adjudicators could not assess their outcome. For example, venography could not be done for a number of participants
Data not available	Participants who are still part of the RCT, however due to incomplete or missing record, some of the outcome data of this participant are missing
Ineligible or mistakenly randomized	Participants who, subsequent to randomization, are either found not to have the condition of interest (e.g. are not pregnant in an RCT among pregnant women), or did not undergo a procedure for which the intervention is intended (e.g. did not undergo surgery in an RCT of postoperative thromboprophylaxis)
Did not receive first dose/treatment	Participants who did not receive the 'first dose' of the intervention to which they were randomized
Ineligible due to early occurrence of outcome	Participants who were eligible at baseline then developed the outcome of interest soon after enrollment. These are considered ineligible if the trialists judge that the occurrence of the outcome cannot be related to the intervention of interest

Experienced adverse events	Participants who developed adverse events but without clear indication whether or not they discontinued the RCT
Non-compliant	Participants who were non-adherent or otherwise violated the protocol
Cross-over	Participants randomized to one arm, but who received the intervention meant for another treatment arm
Withdrew consent	Participants who withdraw their consent to participate in the RCT
Discontinued due to adverse events	Participants who discontinued the RCT due to adverse events
Discontinued trial prematurely	Participants who left the RCT but for whom a reason for discontinuation was not provided
Withdrawn by investigator/clinician	Participants who left the RCT through a decision made by the investigator or clinician (e.g., due to medical necessity)
Unintended protocol violation	Participants who left the RCT due a protocol violation for which they are not responsible (e.g., unavailability of hospital beds)
Lack of efficacy	Participants who left the RCT because they perceived no benefits from the intervention they were randomized to
Protocol violation by investigator/clinician	Investigator/clinician violated the protocol (e.g., change the intended intervention) due to a medical reason
More than one category reported together	The number refers to participants belonging to two or more of the above categories
Other	Reason different than the above

1. Kahale, L.A., et al., *A guidance was developed to identify participants with missing outcome data in randomized controlled trials.* Journal of Clinical Epidemiology, 2019.

Supplementary statistical notes: Data analysis

Under the best-case scenario, Stata imputes missing data as ones in the intervention group and zeroes in the control group. Under the worst-case scenario, Stata imputes missing data as zeroes in the intervention group and ones in the control group. However, the best-case scenario is intended to shift the original effect estimate away from the null value of one, whereas the worst-case scenario is intended to shift the original effect estimate closer to the null value of one. Thus, when applying the worst-case scenario for an outcome with an effect estimate less than 1, we imputed missing data as zeroes in the intervention group and ones in the control group.

When the statistical method of the original meta-analysis of interest was not reported, we used Mantel-Haenszel.

For the calculation of the change in the relative effect estimate, we initially attempted to compare the 'sensitivity analysis pooled relative effect' to the 'original pooled relative effect'. However for the following two reasons, this was not feasible. First, for a significant number of systematic reviews, we could not reproduce the original meta-analysis as it was not clear how the systematic review authors dealt with missing data. Indeed, when we compared the 'best-case scenario pooled relative effect' to the 'original pooled relative effect', 10% of the meta-analyses shifted closer to the null value of one which contradicts the nature of this assumption (i.e., best-case scenario shifts the effect estimate away from the null value of one). Whereas, when we compared the 'best-case scenario pooled relative effect' to the 'complete case analysis pooled relative effect', all meta-analysis shifted away from the null value of one. A very likely explanation is that in first scenario (comparing the 'best-case scenario pooled relative effect' to the 'original pooled relative effect'), missing data were identified and handled differently in the 'original pooled relative effect' by the systematic review authors than how we identified and handled missing data when calculating the 'best-case scenario pooled relative effect'. Second, our approach complies with the GRADE guidance that recommends conducting complete case analysis in the primary analysis and some form of sensitivity analysis, in order to assess the risk of bias associated with missing data (2).

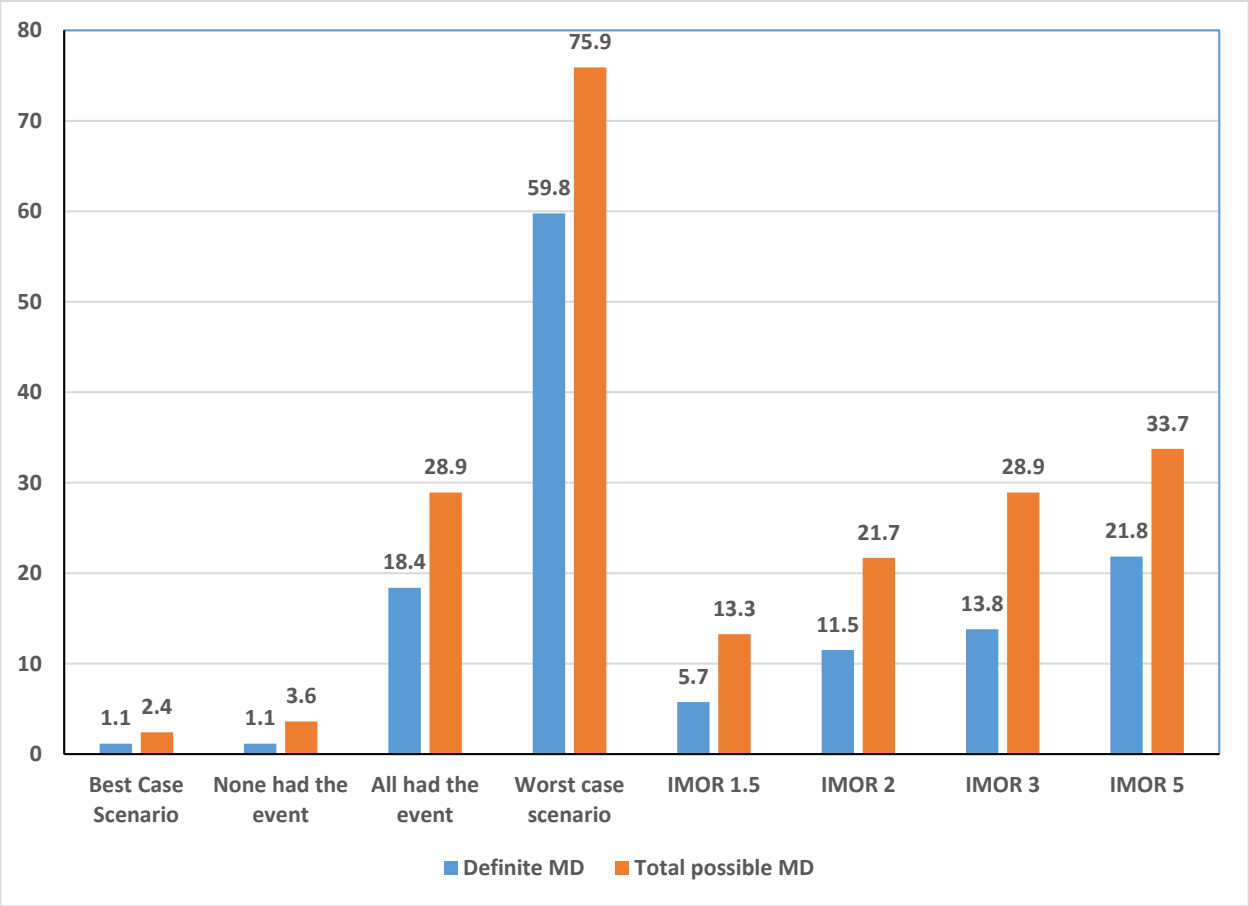
2. Guyatt, G.H., et al., *GRADE guidelines 17: Assessing the Risk of Bias Associated with Missing Participant Outcome Data in a body of evidence.* J Clin Epidemiol, 2017.

Supplementary results 1: The percentage of meta-analyses for which the 'sensitivity analysis pooled relative effect' (1) lost statistical significance and (2) changed direction compared to the 'sensitivity analysis pooled relative effect (CCA)' when considering total possible missing data.

For the four implausible but commonly used assumptions, the percentage of meta-analyses that lost significance varied from 2% (best-case scenario) to 4% (none of the participants with missing data had the event) to 30% (all participants with missing data had the event) to 76% (worst-case scenario). For the plausible assumptions based on IMOR, the percentage of meta-analyses that lost significance varied from 5% (least stringent assumption IMOR 1.5) to 34% (most stringent assumption IMOR 5).

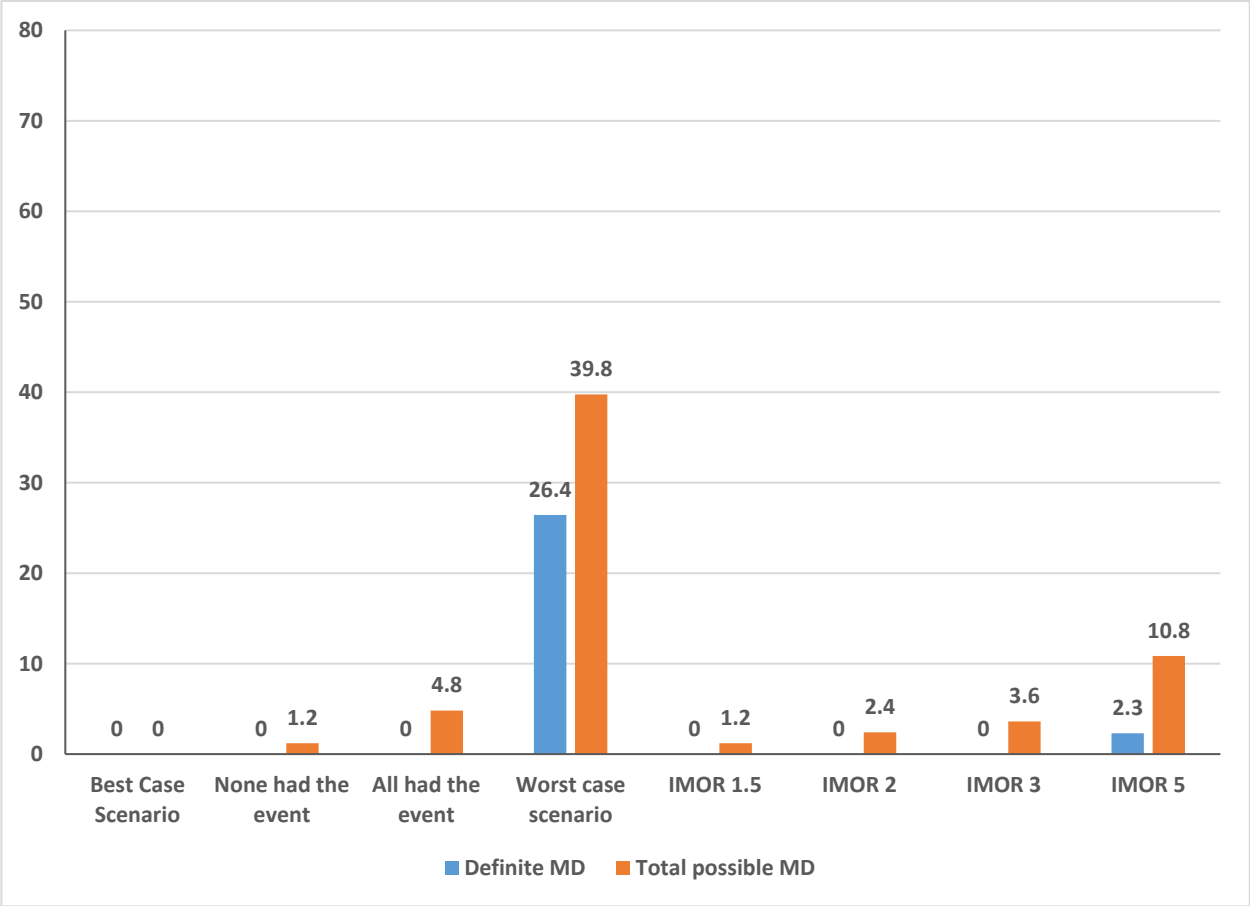
The percentage of meta-analyses that changed direction varied from 0% (best-case scenario), to 1% (none of the participants with missing data had the event), to 5% (all participants with missing data had the event), to 40% (worst-case scenario). As for the five plausible assumptions, the percentage of meta-analyses that changed direction varied from 1.2% (least stringent assumption IMOR 1.5) to 11% (most stringent assumption IMOR 5).

Supplementary results 1 Figure 1: Results of meta-analyses that lost significant when considering participants with definite (in blue) and total possible missing data (in orange) and comparing the 'sensitivity analysis pooled relative effect (assumption)' to the 'sensitivity analysis pooled relative effect (CCA)' (n=87 systematic reviews that maintained statistical significance under the CCA)



***Abbreviations**: IMOR: informative missing odds ratio; MD: missing data*

Supplementary results 1 Figure 2: Results of meta-analyses that changed direction when considering participants with definite (in blue) and total possible missing data (in orange) and comparing the 'sensitivity analysis pooled relative effect (assumption)' to the 'sensitivity analysis pooled relative effect (CCA)' (n=87 systematic reviews that maintained statistical significance under the CCA)



***Abbreviations**: IMOR: informative missing odds ratio; MD: missing data*

Supplementary results 2: The percentage of meta-analyses for which the 'sensitivity analysis pooled relative effect' (1) lost statistical significance and (2) changed direction compared to the 'original pooled relative effect' when considering definite and total possible missing data.

Figures 1 and 2 show the results for the comparison of the 'sensitivity analysis pooled relative effect' to the 'original pooled relative effect' for each method. Specifically, they show the numbers of meta-analyses that lost significant and changed direction respectively, when considering participants with definite and total possible missing data.

Using definite missing data:

Under CCA, the results of 87% of meta-analyses remained statistically significant. For the four implausible but commonly used assumptions, the percentage of meta-analyses that lost significance varied from 3% (best-case scenario) to 12% (none of the participants with missing data had the event) to 27% (all participants with missing data had the event) to 65% (worst-case scenario). For the plausible assumptions based on IMOR, the percentage of meta-analyses that lost significance varied from 18% (least stringent assumption IMOR 1.5) to 32% (most stringent assumption IMOR 5).

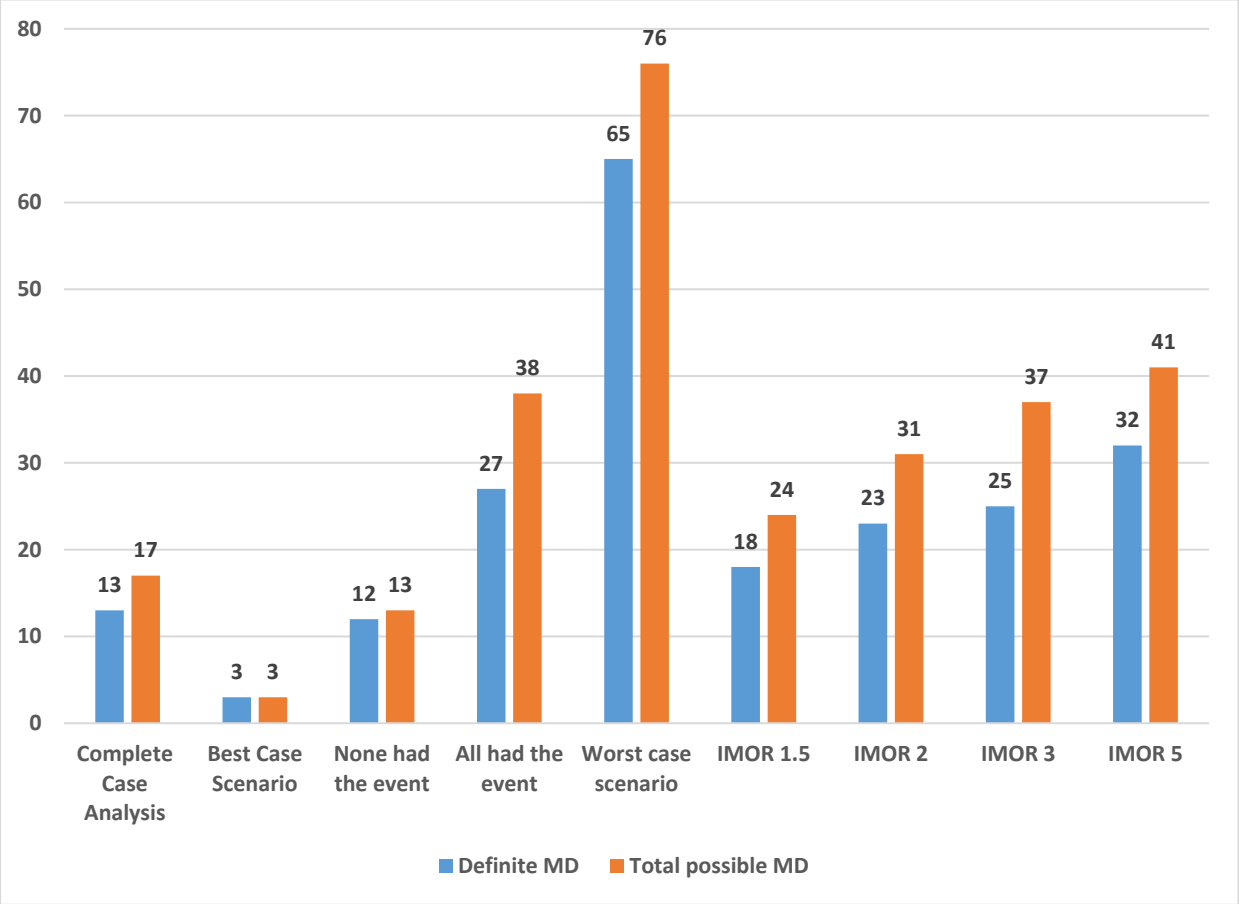
The percentage of meta-analyses that changed direction was 3% under CCA. It varied from 1% (best-case scenario and none of the participants with missing data had the event) to 4% (all participants with missing data had the event) to 33% (worst-case scenario). As for the five plausible assumptions, the percentage of meta-analyses that changed direction varied from 3% (least stringent assumption IMOR 1.5) to 6% (most stringent assumption IMOR 5).

Using total possible missing data:

Under CCA, the results of 83 meta-analyses remained statistically significant. For the four implausible but commonly used assumptions, the percentage of meta-analyses that lost significance varied from 3% (best-case scenario) to 13% (none of the participants with missing data had the event) to 38% (all participants with missing data had the event) to 76% (worst-case scenario). For the plausible assumptions based on IMOR, the percentage of meta-analyses that lost significance varied from 24% (least stringent assumption IMOR 1.5) to 41% (most stringent assumption IMOR 5).

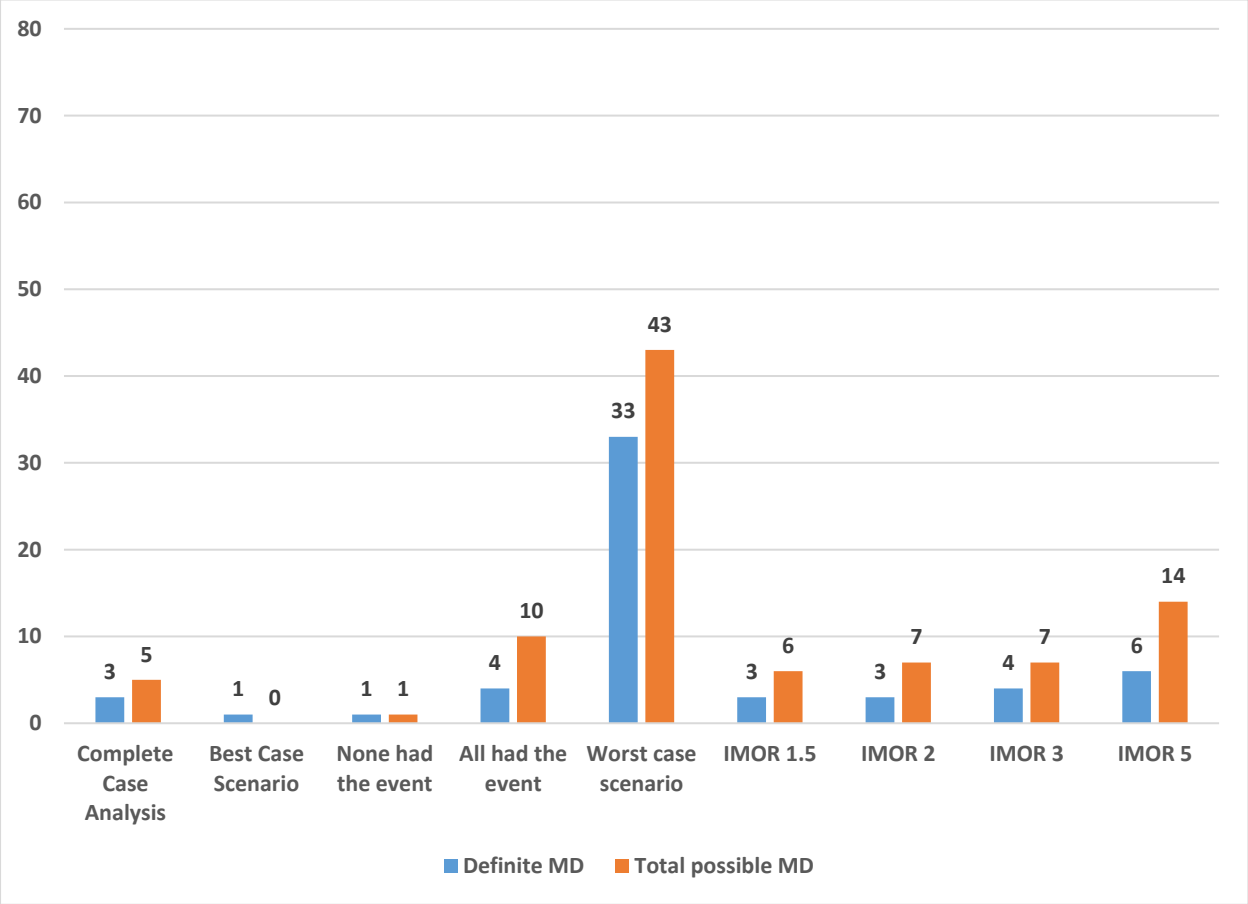
The percentage of meta-analyses that changed direction was 5% under CCA. It varied from 0% (best-case scenario), to 1% (none of the participants with missing data had the event), to 10% (all participants with missing data had the event), and to 43% (worst-case scenario). As for the five plausible assumptions, the percentage of meta-analyses that changed direction varied from 6% (least stringent assumption IMOR 1.5) to 14% (most stringent assumption IMOR 5).

Supplementary results 2 Figure 1: Results of meta-analyses that lost significant when considering participants with definite (in blue) and total possible missing data (in orange) when comparing the 'sensitivity analysis pooled relative effect' to the 'original pooled relative effect' (N=100)



Abbreviations: *IMOR: informative missing odds ratio; MD: missing data*

Supplementary results 2 Figure 2: Results of meta-analyses that changed direction when considering participants with definite (in blue) and total possible missing data (in orange) when comparing the 'sensitivity analysis pooled relative effect' to the 'original pooled relative effect' (n=100)



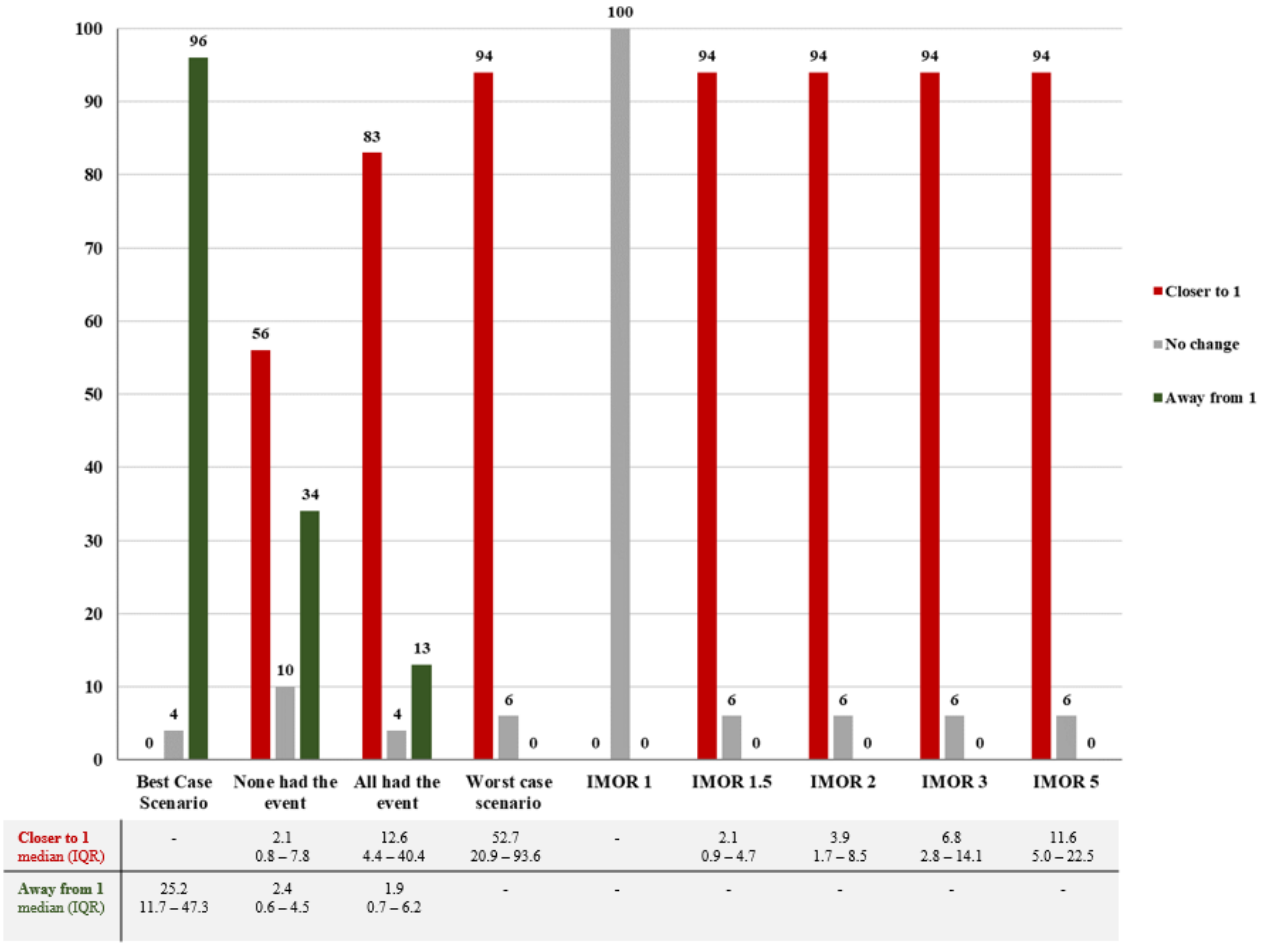
Abbreviations: *IMOR: informative missing odds ratio; MD: missing data*

Supplementary results 3: The percentage change in the relative effect estimate between the 'sensitivity analysis pooled effect estimate (assumption)' and the 'sensitivity analysis pooled effect estimate (CCA)', when considering participants with total possible missing data.

For the four implausible but commonly used assumptions, the percentage of meta-analyses with increased relative effect estimate (shifted away from the null value of 1) was 96% for 'best case scenario' assumption, 34% with 'none of the participants with missing data had the event' assumption, and 13% with 'all participants with missing data had the event' assumption. The median increase in the relative effect estimate varied from 0% for the' worst case scenario' assumption to 25.2% (IQR 11.7%-47.3%) for the 'best case scenario' assumption. The percentage of meta-analyses with reduced relative effect estimate (shifted closer towards the null value of 1) was 94% for the 'worst case scenario' assumption, 56% for 'none of the participants with missing data had the event' assumption, and 83% for 'all participants with missing data had the event' assumption. The median reduction in the relative effect estimate varied from 0% for the 'best case scenario' assumption to 52.8% (IQR 21.5%-94.2%) for the 'worst case scenario' assumption.

For the plausible assumptions based on the IMOR, the percentage of meta-analyses with increased relative effect estimate was 94% of across all stringent assumptions. The median reduction in relative effect estimate varied from 2.1% (IQR 0.9%-4.7%) for IMOR 1.5 assumption to 11.6% (IQR 5.0%-22.5%) for IMOR 5 assumption.

Supplementary results 3 Figure 3: Change of relative effect estimate (by direction) between the 'sensitivity analysis pooled effect estimate (assumption) and the 'sensitivity analysis pooled effect estimate (CCA)' when considering participants with total possible missing data. Bars in the upper part of the figure represent the percentage of meta-analyses with change of relative effect estimate (by direction). The numerical values in the bottom part represent the median (IQR) for, respectively, the increase and decrease in relative effect estimate (N=100)



	Best Case Scenario	None had the event	All had the event	Worst case scenario	IMOR 1	IMOR 1.5	IMOR 2	IMOR 3	IMOR 5
Closer to 1 median (IQR)	-	2.1 0.8 – 7.8	12.6 4.4 – 40.4	52.7 20.9 – 93.6	-	2.1 0.9 – 4.7	3.9 1.7 – 8.5	6.8 2.8 – 14.1	11.6 5.0 – 22.5
Away from 1 median (IQR)	25.2 11.7 – 47.3	2.4 0.6 – 4.5	1.9 0.7 – 6.2	-	-	-	-	-	-

Abbreviations: *IMOR: informative missing odds ratio; IQR: interquartile range; MD: missing data*

Supplementary table 3: Details of the percentage change in the relative effect estimate, stratified by whether the estimate is less than or greater than 1 under the complete case analysis (CCA), using either definite missing data or total possible missing data

Definite missing data

	Best Case Scenario	None had the event	IMOR 1.5	IMOR 2	IMOR 3	IMOR 5	All had the event	Worst case scenario
Effect estimate < 1 under the complete case analysis								
Closer to 1 n (%)	0	24	59	62	62	62	53	63
Median (IQR)	-	4.3 (1.8 – 7.8)	1.3 (0.6 – 2.7)	2.2 (1.0 – 4.8)	4.0 (1.7 – 8.6)	6.6 (2.7 – 14.5)	14.8 (2.8 – 38.2)	30.3 (11.4 – 89.2)
No change n (%)	8	32	13	10	10	10	7	9
Away from 1 n (%)	64	16	0	0	0	0	12	0
Median (IQR)	16.7 (6.1 – 31.5)	1.5 (0.8 – 3.3)	-	-	-	-	0.9 (0.3 – 2.1)	-
Effect estimate > 1 under the complete case analysis								
Closer to 1 n (%)	0	14	26	26	26	26	22	27
Median (IQR)	-	1.3 (0.7 – 4.6)	2.1 (0.7 – 5.0)	3.9 (1.3 – 8.3)	6.5 (2.1 – 12.9)	9.7 (3.2 – 20.6)	10.8 (3.5 – 41.3)	32.6 (10.3 – 65.3)
No change n (%)	1	5	2	2	2	2	1	1
Away from 1 n (%)	27	9	0	0	0	0	5	0
Median (IQR)	22.2 (9.3 – 117.2)	2.5 (1.0 – 9.9)	-	-	-	-	1.9 (0.3 – 3.1)	-

Total possible missing data

	Best Case Scenario	None had the event	IMOR 1.5	IMOR 2	IMOR 3	IMOR 5	All had the event	Worst case scenario
Effect estimate < 1 under the complete case analysis								
Closer to 1 n (%)	0	36	68	68	68	68	62	67
Median (IQR)	-	1.9 (0.7 – 7.8)	1.8 (0.9 – 4.1)	3.6 (1.7 – 7.4)	6.6 (3.2 – 13.9)	9.9 (5.2 – 21.9)	14.7 (4.5 – 39.9)	66.4 (28.9 – 125.8)
No change n (%)	3	9	4	4	4	4	3	5
Away from 1 n (%)	69	27	0	0	0	0	7	0
Median (IQR)	24.6 (11.6 – 39.2)	2.2 (0.7 – 3.7)	-	-	-	-	1.2 (0.9 – 4.7)	-
Effect estimate > 1 under the complete case analysis								
Closer to 1 n (%)	0	20	26	26	26	26	21	27
Median (IQR)	-	2.5 (0.9 – 9.0)	3.4 (1.0 – 5.4)	6.1 (1.7 – 9.7)	9.4 (2.6 – 15.2)	13.6 (3.7 – 22.1)	11.2 (4.4 – 44.6)	35.1 (15.2 – 70.3)
No change n (%)	1	1	2	2	2	2	1	1
Away from 1 n (%)	27	7	0	0	0	0	6	0
Median (IQR)	31.7 (12.4 – 134.6)	4.6 (1.1 – 9.5)	-	-	-	-	2.4 (1.0 – 5.4)	-

References

1. National Research Council, *The Prevention and Treatment of Missing Data in Clinical Trials*. 2010, The National Academic Press: Washington (DC).
2. Akl, E.A., 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**: p. e2809.
3. Bell, M.L., et al., *Handling missing data in RCTs; a review of the top medical journals.* BMC Med Res Methodol, 2014. **14**: p. 118.
4. Fielding, S., et al., *Reporting and dealing with missing quality of life data in RCTs: has the picture changed in the last decade?* Qual Life Res, 2016. **25**(12): p. 2977-2983.
5. Hussain, J.A., et al., *Quality of missing data reporting and handling in palliative care trials demonstrates that further development of the CONSORT statement is required: a systematic review.* J Clin Epidemiol, 2017. **88**: p. 81-91.
6. Powney, M., et al., *A review of the handling of missing longitudinal outcome data in clinical trials.* Trials, 2014. **15**(1): p. 237.
7. Ibrahim, F., et al., *A systematic review of randomised controlled trials in rheumatoid arthritis: the reporting and handling of missing data in composite outcomes.* Trials, 2016. **17**(1): p. 272.
8. Kahale, L.A., et al., *Potentially missing data was considerably more frequent than definitely missing data in randomized controlled trials: A methodological survey.* Journal of Clinical Epidemiology, 2018.
9. Hallgren, K.A., et al., *Missing data in alcohol clinical trials with binary outcomes.* Alcoholism: Clinical and Experimental Research, 2016. **40**(7): p. 1548-1557.
10. Guyatt, G.H., 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.
11. Higgins, J. and S. Green, *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from* www.handbook.cochrane.org. 2011.
12. Akl, E.A., 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**: p. 137.
13. Kahale LA, D.B., Brignardello-Petersen R, Agarwal A, Mustafa RA, Kwong J, Neumann I, Li L, Lopes LC, Briel M, Busse JW, Iorio A, Vandvik PO, Alexander PE, Guyatt G, Akl EA, *Systematic reviews do not adequately report, or address missing outcome data in their analyses: a methodological survey.* Journal of Clinical Epidemiology, 2018.
14. Kahale, L.A., et al., *A guidance was developed to identify participants with missing outcome data in randomized controlled trials.* Journal of Clinical Epidemiology, 2019.
15. Harris, P.A., et al., *Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support.* J Biomed Inform, 2009. **42**(2): p. 377-81.
16. SPSS, *IBM SPSS Statistics for Windows. Version 21.0.* 2013, IBM Corp: Armonk: NY.
17. StataCorp, *Stata Statistical Software: Release 12*, C. Station, Editor. 2011, StataCorp LP: TX.
18. Akl, E.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**: p. 98.
19. White, I.R., et al., *Allowing for uncertainty due to missing data in meta-analysis--part 2: hierarchical models.* Stat Med, 2008. **27**(5): p. 728-45.

20. White, I.R., J.P. Higgins, and A.M. Wood, *Allowing for uncertainty due to missing data in meta-analysis--part 1: two-stage methods.* Stat Med, 2008. **27**(5): p. 711-27.
21. White, I.R. and J.P. Higgins, *Meta-analysis with missing data.* The Stata Journal, 2009. **9**(1): p. 57-69.
22. Gamble, C. and S. Hollis, *Uncertainty method improved on best-worst case analysis in a binary meta-analysis.* J Clin Epidemiol, 2005. **58**(6): p. 579-88.
23. Wasserstein, R.L., A.L. Schirm, and N.A. Lazar, *Moving to a World Beyond 'p< 0.05'*. 2019, Taylor & Francis.
24. Ebrahim, S., et al., *Addressing continuous data for participants excluded from trial analysis: a guide for systematic reviewers.* J Clin Epidemiol, 2013. **66**(9): p. 1014-1021 e1.
25. Ebrahim, S., et al., *Addressing continuous data measured with different instruments for participants excluded from trial analysis: a guide for systematic reviewers.* J Clin Epidemiol, 2014. **67**(5): p. 560-70.
26. Sullivan, T.R., et al., *Treatment of missing data in follow-up studies of randomised controlled trials: A systematic review of the literature.* Clin Trials, 2017. **14**(4): p. 387-395.
27. Babic, A., et al., *Assessments of attrition bias in Cochrane systematic reviews are highly inconsistent and thus hindering trial comparability.* BMC Medical Research Methodology, 2019. **19**(1): p. 76.
28. Shivasabesan, G., B. Mitra, and G.M. O'Reilly, *Missing data in trauma registries: A systematic review.* Injury, 2018. **49**(9): p. 1641-1647.
29. Hollestein, L. and J. Carpenter, *Missing data in clinical research: an integrated approach.* British Journal of Dermatology, 2017. **177**(6): p. 1463-1465.
30. Little, R., et al., *The design and conduct of clinical trials to limit missing data.* Statistics in medicine, 2012. **31**(28): p. 3433-3443.
31. Akl, E.A., et al., *Reporting missing participant data in randomised trials: systematic survey of the methodological literature and a proposed guide.* BMJ Open, 2015. **5**(12): p. e008431.
32. Mavridis, D. and I.R. White, *Dealing with missing outcome data in meta-analysis.* Research synthesis methods, 2019.
33. Spineli, L.M., *An empirical comparison of Bayesian modelling strategies for missing binary outcome data in network meta-analysis.* BMC medical research methodology, 2019. **19**(1): p. 86.

Chapter 6

Meta-analyses proved inconsistent in how missing data were handled across their included primary trials: a methodological survey

Kahale LA, Khamis AM, Diab B, Chang Y, Lopes LC, Agarwal A, Li L, Mustafa R, Koujanian S, Waziry R, Busse JW, Dakik A, Schünemann H, Hooft L, Guyatt G, Scholten RJPM, and Akl EA. International Journal of Epidemiology (submitted)

Abstract

Background and Objective:

The extent to which systematic review authors address missing data consistently across eligible primary studies within the same meta-analysis and in a way that is consistent with their reported methods remains uncertain. The objectives were to assess whether the systematic review authors are consistent in the way they handle missing data, both across trials included in the same meta-analysis, and with their reported methods.

Methods:

We identified 100 eligible systematic reviews that included a group-level meta-analysis of a patient-important dichotomous efficacy outcome, with a statistically significant effect estimate and the 653 trials included in these review's meta-analyses. From each trial report, we abstracted statistical data used in the analysis of the outcome of interest and compared the statistical data from the trial report to the data included in the meta-analysis. First, we used these comparisons to classify the 'analytical method actually used' for handling missing data by the systematic review authors for each trial report. Second, we assessed whether systematic reviews explicitly reported on the analytical method of handling missing data. Third, we calculated the proportion of systematic reviews that were consistent in the 'analytical method actually used' across trials included in the same meta-analysis. Fourth, among the systematic reviews that were consistent in the 'analytical method actually used' across trials and explicitly reported on a method for handling missing data, we assessed whether the 'analytical method actually used' and the reported methods were consistent.

Results:

We were not able to classify the 'analytical method reviews actually used' for 397 RCTs. Among the remaining 241, systematic review authors conducted 'complete case analysis' in 128 (53%) and assumed 'none of the participants with missing data had the event of interest' in 58 (24%). Second, we found that only eight out of 100 systematic reviews were consistent in handling missing data across the included trials. Third, seven out of the 100 systematic reviews explicitly reported on the analytical method of handling missing data. Fourth, of these seven systematic reviews, one was consistent in handling

missing data across the included trials (using complete case analysis); however, that method was not consistent with their reported methods.

Conclusion:

We found that systematic review authors were inconsistent in their methods of handling missing data across their eligible primary trials. Moreover, most systematic review authors did not explicitly report their methods to handle missing data. Of the seven reviews that did explicitly report on their methods, none applied that method consistently across the included trials. As such inconsistency might threaten the validity of the results of systematic reviews, methodologic rigor requires improved adherence to guidance on identifying, reporting, and handling participants with missing outcome data.

Introduction

Reporting whether outcome data in randomized controlled trials (RCTs) are missing is often suboptimal [1]. First, instead of reporting missing data information specific to each outcome, RCTs typically report the number of participants with premature end of follow-up in general [2]. However, those who had premature end of follow-up will not necessarily have data missing for all outcomes. For example, they might have experienced certain outcomes (and have them documented) prior to their loss of follow-up. Second, it is not always clear whether RCT authors followed-up certain participants such as those who withdrew consent to be part of the RCT (i.e., whether they have missing data or not) [1]. Third, RCT authors often fail to clearly describe how they dealt with missing data in their analyses (e.g., complete case analysis, imputation for missing data) [1, 2].

The poor reporting of missing outcome data in RCTs necessitates that systematic review authors develop plans to address them [3-16]. However, a recent methodological survey found that only 25% of systematic review authors planned to consider whether certain categories of participants (e.g., withdrew consent, non-compliant) might have missing outcome data [17]. In addition, the survey found that only 19% reported a method to handle missing data (e.g., complete case analysis, making assumptions) [17].

However, when systematic review authors decide to handle missing outcome data in the analysis, they may do so inconsistently across trials included in the same meta-analysis. As an illustrative scenario, a systematic review reports in the methods section a plan to include only participants with available outcome data in their meta-analysis (i.e., use complete case analysis). One would then expect the denominators of all trials included in that meta-analysis to be restricted to only participants with available outcome data. However, very often, for one trial reviewer authors may use the total number randomized for the denominator (despite having participants with missing data) and in another trial they may exclude participants with missing data from the denominator. In such a scenario, we observe two main potential problems: (1) the analytical method review authors actually used for handling missing data is inconsistent across trials included in the same meta-analysis; and (2) the analytical method review authors actually used for handling missing data is for some trials inconsistent with the reported methods. These

two problems might hinder the reproducibility of the systematic reviews and may even bias results. The extent of these problems remains, however, unclear.

Objective

The overall objective of this study was to assess whether the systematic review authors are consistent in the way they handle missing data, both across trials included in the same meta-analysis, and with their reported methods. More specifically, we aimed to: (1) classify the methods systematic review authors actually used for handling missing data for each included trial; (2) assess whether systematic reviews authors explicitly reported on the method of handling missing data; (3) assess the extent to which systematic review authors were consistent in their methods actually used across trials included in the same meta-analysis; (4) when consistent, assess whether the methods the systematic review authors actually used were consistent with their reported methods (if reported).

Methods

Study design and definitions

This methodological study is part of a larger project examining methodological issues related to missing outcome data in systematic reviews and RCTs [18]. Our published protocol includes detailed information on the definitions, eligibility criteria, search strategy, selection process, data abstraction and data analysis [18]. We defined missing data as outcome data for trial participants that are not available to systematic review authors from the published RCT reports or personal contact with RCT authors. We used our recently published guidance [19] to identify categories of trial participants who might have missing outcome data.

Sample selection

Our random sample included a 50 Cochrane and 50 non-Cochrane systematic reviews published in 2012 that reported a group-level meta-analysis of a patient-important dichotomous efficacy outcome, with a statistically significant effect estimate [17]. We retrieved all 653 RCTs included in the 100 meta-analyses of interest [1]. In duplicate and

independently, eleven pairs of reviewers abstracted data from the systematic reviews and RCTs and resolved disagreements with the help of a third reviewer. We conducted calibration exercises and used standardized and pilot-tested forms with detailed written instructions.

Classifying the 'analytical method reviews actually used' for handling missing data

Authors of reviews may fail to clearly report their approach to handling missing data. Alternatively, the approach they state in the methods they use may not correspond with the method they actually used. Therefore, we specified the concept of the 'analytical method reviews actually used' for handling missing data using the following steps:

1. From each RCT report, we abstracted (per study arm) the number of participants randomized, the numerator (i.e. the number of events) used in the analysis of interest, and the number of participants with missing data;
2. From the meta-analysis (forest plot plus text) and for each arm of all contributing RCTs, we abstracted the denominator and the numerator used in the meta-analysis of interest;
3. We compared the statistical data from the RCT report with data from the meta-analysis;
4. Based on this comparison, we classified the 'analytical method reviews actually used' for handling missing data as:
    - Unclear, cannot be verified (provided numbers that could not be explained or did not add up to match any of the suggested analytical method actually used);
    - Complete case analysis;
    - Making assumptions (e.g., best case scenario, all participants had the event);
    - Different methods (from the above bullet points) for different categories of participants with missing data;
    - Not applicable, no missing data.

The hypothetical examples in Table 1 illustrate how a number of possible meta-analyses addressing the same study question (i.e. same patients, interventions, comparators and outcomes) might have dealt with missing data from a single eligible RCT and how we classified the 'analytical method reviews actually used' in each case. We also assessed the

confidence of data abstractors in classifying the analytical method actually used i.e., whether based on explicit reporting (higher confidence) or best guess (lower confidence).

Also, for systematic reviews that reported on having participants with missing outcome data, we assessed whether the systematic review authors used the same denominator and/or numerator as the one reported in the RCT that informed their meta-analysis.

Consistency in analytical methods

After classifying the 'analytical method reviews actually used' for handling missing data (aim 1), first, we assessed whether the authors explicitly reported on the analytical method of handling missing data, which if present we designated as the 'reported analytical method' (aim 2). Second, for each meta-analysis, we assessed whether the 'analytical method reviews actually used' for handling missing data was consistent across trials within this meta-analysis (aim 3). If the answer was yes to both of the above, we addressed whether the 'analytical method reviews actually used' of handling missing data in the meta-analysis was consistent with its 'reported analytical method' (aim 4). We displayed the results of the 'reported' and 'actual' analytical methods in a matrix (see table 2). Table 2 presents hypothetical scenarios illustrating the different judgments made.

Statistical analysis

Using SPSS statistical software, version 21.0 [20], we conducted a descriptive analysis (frequencies and percentages) of all collected variables. We also planned to conduct regression analyses to study the association between 'consistency between actual and reported method' and the characteristics of the included systematic review. However, we did not run these analyses due to the low number of reviews that were consistent within the same meta-analysis.

Results

The sample of 100 eligible systematic reviews with significant pooled effect estimates included 653 RCTs; we were able to get the full texts for 638 RCTs. We previously reported on the details of these systematic reviews [17] and the included RCTs [1].

Classifying the 'analytical method reviews actually used' for handling missing data

Table 3 summarizes the characteristics of the 'actual analytical methods' used by the systematic review authors across all included RCTs. We were able to classify the 'analytical method reviews actually used' for 241 (38%) of the included RCTs. The data abstractors were able to classify the analytical method actually used for 67% of the 241 based on their best guess and for the remaining based on explicit reporting. For the remaining RCTs, 207 (32%) included no participants with missing data (complete follow-up), 161 (25%) provided numbers that could not be explained (did not add up to match any of the suggested analytical method actually used), 5 (1%) had data wrongly abstracted (e.g., entered in the meta-analysis data from the wrong outcome), and 24 (4%) provided insufficient information from both RCT and systematic review reports.

Among the 241 included RCTs for which we were able to classify the 'analytical method reviews actually used', systematic review authors conducted 'complete case analysis' in 128 (53%), assumed 'none of the participants with missing data had the event of interest' in 58 (24%), and used different methods for different categories of participants with missing data in 51 (21%) of RCTs. In four RCTs (2%) various other assumptions were used (Table 3).

For the 'reported analytical method', among 100 systematic reviews, only seven reported on methods to handle missing data in the meta-analysis. Two reported conducting complete case analysis, two reported assuming that all of participants with missing data had the event of interest, and three reported assuming that none of participants with missing data had the event of interest.

When we explored whether the systematic review authors, who reported on having participants with missing outcome data, used in their meta-analysis the same

denominator as the one reported in the RCT, we found that in the majority of cases they used a denominator (81%) and a numerator (80%) reported by the RCT (Table 3).

Consistency in analytical methods within meta-analyses

Of the seven systematic reviews that explicitly reported on the analytical method of handling missing data, only one was consistent in handling missing data across all included trials (using complete case analysis) (Figure 1). However, the analytical method actually used was not consistent with their 'reported analytical methods' ('if missing data were unable to be obtained, a result was assumed to have a particular value, such as poor outcome') [21]. Of the 93 systematic reviews that did not explicitly report their analytical method of handling missing data, seven were consistent in their actual analytical method for handling missing data across all included trials. So, in total eight systematic reviews were consistent in handling missing data across the included trials.

Discussion

Summary of findings

We found that systematic review authors were inconsistent in their methods of handling missing data across their eligible primary trials. Moreover, most systematic review authors did not explicitly report their methods to handle missing data. Of the seven reviews that did explicitly report on their methods, none applied that method consistently across the included trials.

Strengths and limitations

The main strength of our study is the systematic and transparent methods used, including screening independently and in duplicate, conducting calibration exercises, using pilot tested forms for data abstraction, and applying a detailed, carefully constructed, and logically coherent strategy for making the numerous classification judgments involved in the study. Up to our knowledge, this is first methodological survey that explores how systematic review authors actually dealt with trial missing data in their meta-analysis. Also, this is the first study to assess whether the methods used for handling missing outcome data in the meta-analysis are consistent with the 'reported analytical methods'.

One limitation of our study is the reliance on reviewers' judgments at different stages of the process (e.g., judgment about the 'actually analytical methods'). The prior development of classification systems, the specific and detailed instructions, pilot testing, and calibration exercises may mitigate this concern. Further, our sample included systematic reviews that were published in 2012, and may not reflect more current reviews; however, recent surveys have found the reporting, handling, and assessment of risk of bias in relation to missing data has not improved over this period of time [16, 22-24].

Interpretation of findings

Both the challenge we faced in classifying the 'analytical method reviews actually used' (25% of RCTs provided numbers that could not be explained), and the observed inconsistency in handling missing data within the same meta-analysis relate to the failure of reviewers to adopt standardized approaches to reporting and dealing with missing data [19, 25, 26]. This inconsistency might bias the results and could be a reason for variations between the results of different meta-analyses addressing the same research question [27].

We uncovered three limitations in how systematic reviews authors handle missing data in their meta-analysis: lack of transparency in reporting the method of dealing with missing data:

1. 93% did not explicitly report on their methods for handling missing data;
2. 92% were inconsistent in the methods used to handle missing data across RCTs within the same meta-analysis;
3. In the few meta-analyses that did explicitly report the method they proposed to use to handle missing data, none actually used that method.

We also found that for more than 80% of RCTs with missing outcome data contributing to the meta-analyses of interest, the systematic review authors used the same denominator and numerator as those reported by the trialists. So, systematic review authors may simply use what trialists have reported, without consciously planning a method to handle missing data. This practice might explain why systematic review authors are not consistent with their approach in handling missing data across trials

included in the same meta-analysis. In some other cases, systematic review authors and trial authors, and with an intention to apply the 'intention to treat' principle, include in the denominator the total number randomized while using whatever the trial authors have used in the numerator. Subsequently, they would be implicitly applying 'none of the participants with missing data had the event'.

Implications for practice

In order to ensure consistency in handling missing data across trials included in the same meta-analysis, authors should:

1. Develop a transparent and detailed strategy for handling missing data (e.g., using complete case analysis, applying assumptions) [28-32];
2. Refer to available guidance on how to identify participants with missing data from RCT reports [19];
3. Apply their strategy for handling missing data consistently across all trials included in the meta-analysis;
4. Report clearly on the above.

Conclusion

The large majority of systematic reviews considered in our study did not report a method for handling missing data in their meta-analyses. For the few that did, the actual method used for handling missing outcome data was often inconsistent with their reported methods. As such inconsistency might threaten the validity of the results of systematic reviews, methodologic rigor requires improved adherence to guidance on identifying, reporting, and handling participants with missing outcome data.

Figures, tables, and supplementary data

Legends

Table 1: Hypothetical examples illustrating how a number of meta-analyses addressing the same study question might have handled missing data for an unfavorable outcome from a single eligible RCT report and thus informed classification of the 'analytical method reviews actually used'.

Table 2: Hypothetical scenarios illustrating the process for judging consistency between 'reported' and 'actual' analytical methods for addressing missing data

Table 3: Characteristics of the 'analytical method reviews actually used' used by the systematic review authors across all included RCTs

Figure 1: Flow diagram showing consistency in analytical methods within the same meta-analysis and versus the reported analytical method

Table 1: Hypothetical examples illustrating how a number of meta-analyses addressing the same study question might have handled missing data for an unfavorable outcome from a single eligible RCT report and thus informed classification of the 'analytical method reviews actually used'

	Intervention arm			Control arm		
RCT report data						
	Number of participants randomized	Number of events	Number of participants with missing data	Number of participants randomized	Number of events	Number of participants with missing data
RCT 1	100	5	2	100	10	5
Meta-analysis data						
	Denominator		Numerator	Denominator	Numerator	Our classification of the actual analytical method by the SR
MA 1	98		5	95	10	Complete case analysis
MA 2	100		5	100	10	Assumed none had the event
MA 3	100		7	100	15	Assumed that had the event
MA 4	100		7	100	10	Worst-case scenario
MA 5	100		5	100	15	Best-case scenario

Abbreviations: *MA: meta-analysis; RCT: randomized controlled trial; SR: systematic review*

Table 2: Hypothetical scenarios illustrating the process for judging consistency between 'reported' and 'actual' analytical methods for addressing missing data

		Reported analytical method	Actual analytical method	Actual analytical method consistent within the meta-analysis	Actual analytical method consistent with reported analytical method
SR 1		Assume all had the event			
	RCT 1	-	Complete case analysis	No	Not applicable since the actual analytical method were inconsistent across trials
	RCT 2	-	Assume none had the event		
	RCT 3	-	Different methods for different categories of participants with missing data		
SR 2		Complete case analysis			
	RCT 4	-	Complete case analysis	Yes	Yes
	RCT 5	-	Complete case analysis		
	RCT 6	-	Complete case analysis		
SR 3		Assume all had the event			
	RCT 7	-	Complete case analysis	Yes	No
	RCT 8	-	Complete case analysis		
	RCT 9	-	Complete case analysis		
SR 4		Not reported			
	RCT 10	-	Complete case analysis	Yes	Not applicable since the reported analytical method is not available
	RCT 11	-	Complete case analysis		
	RCT 12	-	Complete case analysis		
SR 5		Not reported			
	RCT 13	-	Complete case analysis	No	Not applicable since the reported analytical method is not available
	RCT 14	-	Assume none had the event		
	RCT 15	-	Different methods for different categories of participants with missing data		

Table 3: Characteristics of the 'analytical method reviews actually used' used by the systematic review authors across all included RCTs

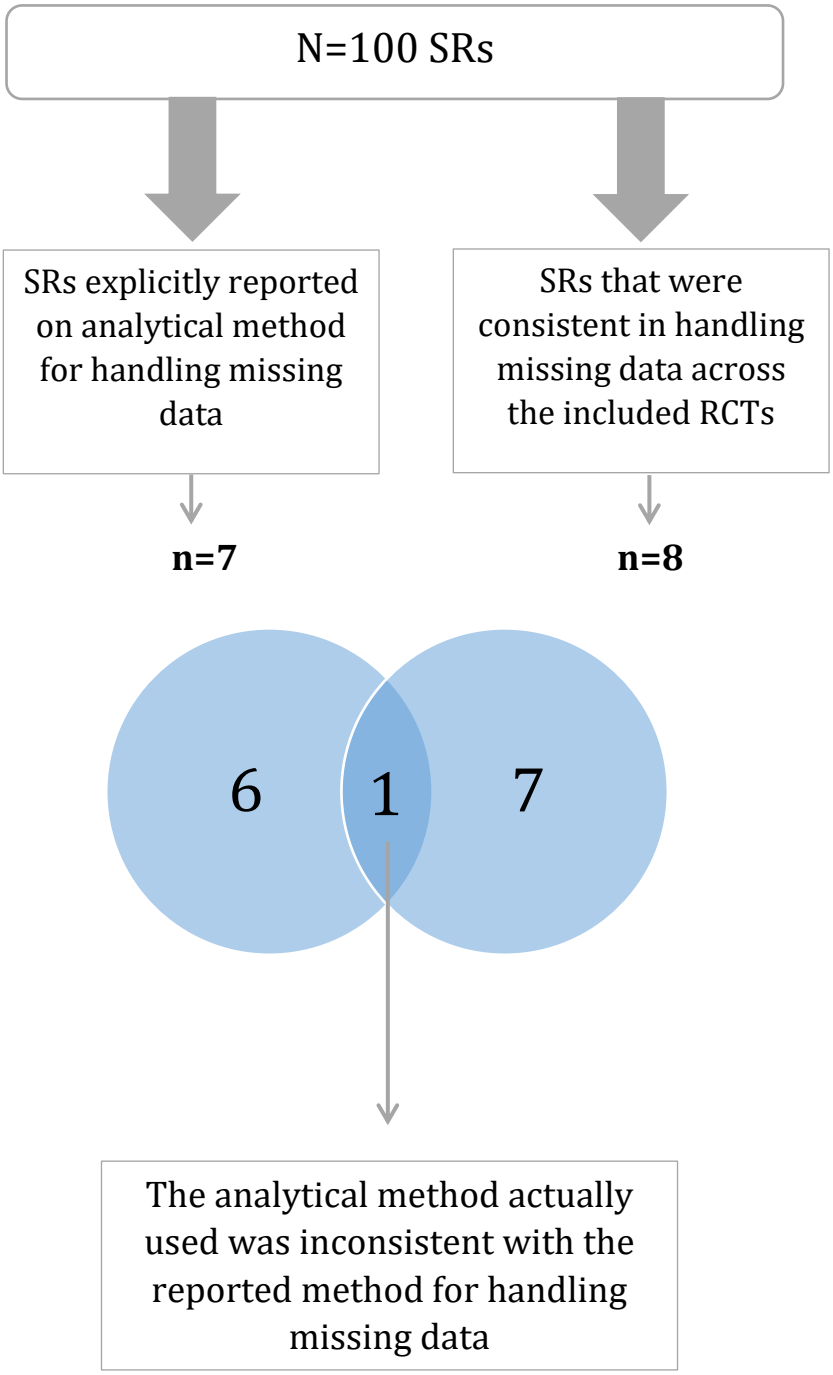
Variable	n (%)
Ability to classify the *'analytical method reviews actually used'* (n=638)	
Able to classify	241 (37.8)
Not applicable (no missing data)	207 (32.4)
Could not be explained (numbers do not add up)	161 (25.2)
Wrong data abstraction	5 (0.8)
No data available from RCT or SR	24 (3.8)
Classification of the *'analytical method reviews actually used'* (n= 241[+])	
Complete case analysis	128 (53.1)
Assumption: none of the participants with missing data had the event of interest	58 (24.1)
Assumption: 'all of the participants with missing data had the event of interest'	2 (0.8)
Assumption: worst-case scenario	1 (0.4)
Assumption: best-case scenario	0
Assumption: same event rate as those followed up	0
Other assumption	1 (0.4)
Different methods for different categories of participants with missing data	51 (21.2)
The SR authors used in the meta-analysis a denominator used by the RCT (n=431*)	
Definitely yes	348 (80.7)
Definitely no	65 (15.1)
Unclear	18 (4.2)
The SR authors used in the meta-analysis a numerator used by the RCT (n=431*)	
Definitely yes	345 (80.0)
Definitely no	53 (12.3)
Unclear	33 (7.7)

Abbreviations: *ITT: intention-to-treat; LTFU: lost to follow-up; RCT: randomized controlled trials; SR: systematic reviews*

[+]n=241 RCTs for which we could classify an actual analytical method

*n= 128 RCTs for which the SR authors actually conducted a complete case analysis

Figure 1: Flow diagram showing consistency in analytical methods within the same meta-analysis and versus the reported analytical method



Abbreviations: *RCTs: randomized controlled trials; SR: systematic review*

References

1. Kahale LA, Diab B, Khamis AM, et al. Potentially missing data was considerably more frequent than definitely missing data in randomized controlled trials: A methodological survey. Journal of Clinical Epidemiology 2018
2. Akl EA, Kahale LA, Ebrahim S, et al. Three challenges described for identifying participants with missing data in trials reports, and potential solutions suggested to systematic reviewers. J Clin Epidemiol 2016 doi: 10.1016/j.jclinepi.2016.02.022
3. Hussain JA, White IR, Langan D, et al. Missing data in randomised controlled trials evaluating palliative interventions: a systematic review and meta-analysis. The Lancet 2016;387:S53. doi: 10.1016/s0140-6736(16)00440-2
4. Adewuyi TE, MacLennan G, Cook JA. Non-compliance with randomised allocation and missing outcome data in randomised controlled trials evaluating surgical interventions: a systematic review. BMC Res Notes 2015;8:403. doi: 10.1186/s13104-015-1364-9
5. Akl EA, Shawwa K, Kahale LA, et al. Reporting missing participant data in randomised trials: systematic survey of the methodological literature and a proposed guide. BMJ Open 2015;5(12):e008431. doi: 10.1136/bmjopen-2015-008431
6. Bell ML, Fiero M, Horton NJ, et al. Handling missing data in RCTs; a review of the top medical journals. BMC Med Res Methodol 2014;14:118. doi: 10.1186/1471-2288-14-118
7. Fielding S, Maclennan G, Cook JA, et al. A review of RCTs in four medical journals to assess the use of imputation to overcome missing data in quality of life outcomes. Trials 2008;9:51. doi: 10.1186/1745-6215-9-51
8. Fielding S, Ogbuagu A, Sivasubramaniam S, et al. Reporting and dealing with missing quality of life data in RCTs: has the picture changed in the last decade? Qual Life Res 2016;25(12):2977-83. doi: 10.1007/s11136-016-1411-6
9. Fiero MH, Huang S, Oren E, et al. Statistical analysis and handling of missing data in cluster randomized trials: a systematic review. Trials 2016;17:72. doi: 10.1186/s13063-016-1201-z
10. Gewandter JS, McDermott MP, McKeown A, et al. Reporting of missing data and methods used to accommodate them in recent analgesic clinical trials: ACTTION systematic review and recommendations. Pain 2014;155(9):1871-7. doi: 10.1016/j.pain.2014.06.018
11. Karlson CW, Rapoff MA. Attrition in randomized controlled trials for pediatric chronic conditions. J Pediatr Psychol 2009;34(7):782-93. doi: 10.1093/jpepsy/jsn122
12. Masconi KL, Matsha TE, Echouffo-Tcheugui JB, et al. Reporting and handling of missing data in predictive research for prevalent undiagnosed type 2 diabetes mellitus: a systematic review. EPMA J 2015;6(1):7. doi: 10.1186/s13167-015-0028-0
13. Spineli LM. Missing binary data extraction challenges from Cochrane reviews in mental health and Campbell reviews with implications for empirical research. Res Synth Methods 2017;8(4):514-25. doi: 10.1002/jrsm.1268
14. Wahlbeck K, Tuunainen A, Ahokas A, et al. Dropout rates in randomised antipsychotic drug trials. Psychopharmacology 2001;155(3):230-33. doi: 10.1007/s002130100711
15. Wood AM, White IR, Thompson SG. Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. Clin Trials 2004;1(4):368-76.

16. Ibrahim F, Tom BD, Scott DL, et al. A systematic review of randomised controlled trials in rheumatoid arthritis: the reporting and handling of missing data in composite outcomes. Trials 2016;17(1):272.
17. Kahale LA DB, Brignardello-Petersen R, Agarwal A, Mustafa RA, Kwong J, Neumann I, Li L, Lopes LC, Briel M, Busse JW, Iorio A, Vandvik PO, Alexander PE, Guyatt G, Akl EA. Systematic reviews do not adequately report, or address missing outcome data in their analyses: a methodological survey. Journal of Clinical Epidemiology 2018 doi: 10.1016/j.jclinepi.2018.02.016
18. Akl EA, Kahale LA, Agarwal A, 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. doi: 10.1186/2046-4053-3-137
19. Kahale LA, Guyatt GH, Agoritsas T, et al. A guidance was developed to identify participants with missing outcome data in randomized controlled trials. Journal of Clinical Epidemiology 2019
20. IBM SPSS Statistics for Windows. Version 21.0. [program]. Armonk: NY: IBM Corp, 2013.
21. Tang H, Hunter T, Hu Y, et al. Cabergoline for preventing ovarian hyperstimulation syndrome. Cochrane Database of Systematic Reviews 2012(2)
22. Sullivan TR, Yelland LN, Lee KJ, et al. Treatment of missing data in follow-up studies of randomised controlled trials: A systematic review of the literature. Clin Trials 2017;14(4):387-95. doi: 10.1177/1740774517703319
23. Babic A, Tokalic R, Cunha JAS, et al. Assessments of attrition bias in Cochrane systematic reviews are highly inconsistent and thus hindering trial comparability. BMC Medical Research Methodology 2019;19(1):76.
24. Shivasabesan G, Mitra B, O'Reilly GM. Missing data in trauma registries: A systematic review. Injury 2018;49(9):1641-47. doi: 10.1016/j.injury.2018.03.035
25. Akl EA, Carrasco-Labra A, Brignardello-Petersen R, 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. doi: 10.1136/bmjopen-2015-009368
26. Akl EA, Kahale LA, Agoritsas T, 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. doi: 10.1186/s13643-015-0083-6
27. Khamis AM, El Moheb M, Nicolas J, et al. Several reasons explained the variation in the results of 22 meta-analyses addressing the same question. J Clin Epidemiol 2019;113:147-58. doi: 10.1016/j.jclinepi.2019.05.023
28. Akl EA, Johnston BC, Alonso-Coello P, et al. Addressing dichotomous data for participants excluded from trial analysis: a guide for systematic reviewers. PLoS One 2013;8(2):e57132. doi: 10.1371/journal.pone.0057132
29. Guyatt GH, Ebrahim S, Alonso-Coello P, 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 doi: 10.1016/j.jclinepi.2017.05.005
30. White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. Clin Trials 2012;9(4):396-407. doi: 10.1177/1740774512450098
31. White IR, Higgins JP. Meta-analysis with missing data. The Stata Journal 2009;9(1):57-69.

32. White IR, Horton NJ, Carpenter J, et al. Strategy for intention to treat analysis in randomised trials with missing outcome data. BMJ 2011;342:d40. doi: 10.1136/bmj.d40

Chapter 7

General Discussion

Introduction

Missing outcome data of trial participants is a frequent phenomenon in RCTs and may represent a serious potential source of bias if not reported and handled appropriately. The potential effect of bias associated with missing outcome data - attrition bias - is that invalid conclusions about efficacy and safety of studied interventions may be reached and ultimately impact clinical practice.

Certain categories of participants that trialists report on (e.g., lost to follow-up, withdrew consent, non-compliant) might have missing outcome data. Trialists cease following-up certain categories of participants (e.g., lost to follow-up), thus these categories would definitely have missing data. On the other hand, it is not clear whether trialists follow-up other categories of participants (e.g., non-compliant), thus these categories would potentially have missing data. In some cases, these categories provide the only information about missing outcome data that systematic review authors can use in their meta-analyses. The poor reporting and handling of missing outcome data in RCTs contribute to the inadequate reporting and handling of missing outcome data in systematic reviews. Systematic review authors need specific guidance on how to address missing outcome data of trial participants in their reviews. The overall aim of this thesis was to address this need.

In this chapter, we present a summary of our results. We then discuss the strengths and limitations of our research and finally, we provide the implications of our findings for various audiences.

Summary

In **chapter 2**, we described how systematic review authors report on participants who might have missing outcome data, handle missing data in their primary meta-analyses of dichotomous outcomes, and assess the associated risk of bias. The key findings of chapter 2 were:

- We defined a list of 10 categories of trial participants who might have missing outcome data.
- Systematic review authors do not explicitly report sufficient information on categories of trial participants who might have missing outcome data, which may have implications for addressing those in the meta-analysis.
- The most reported category among systematic review authors is ‘‘unexplained loss to follow-up’’.
- Systematic review authors do not typically judge risk of bias associated with missing outcome data at the level of the meta-analysis.
- Systematic review authors do not typically handle missing outcome data in their primary analyses.
- Complete case analysis is the most frequently used method by systematic review authors to handle missing outcome data in their primary meta-analyses.

In **chapter 3**, we assessed how trial authors report on the categories of participants that might have missing outcome data and their follow-up status, and on the handling of these participants in their main and secondary analyses. The key findings of chapter 3 were:

- We refined the original list of pre-defined categories of participants that might have missing outcome data to accommodate new categories that emerged from data abstraction and did not fit existing categories. The refined list included 19 categories.
- In RCTs, potentially missing outcome data is considerably more frequent than definitely missing outcome data.
- None of the RCTs that imputed outcomes, took into account the uncertainty associated with imputing outcomes.
- None of the RCTs reported on participants that might have experienced certain outcomes (and have them documented) prior to their loss of follow-up.
- When studies explicitly reported not following-up participants with missing data, about half explicitly reported conducting a complete case analysis; almost all of the remainder studies did not specify how they handled missing outcome data in their analysis.

In **chapter 4**, we developed guidance for authors of systematic reviews on how to identify and classify participants with missing outcome data in trials so that they can better investigate the impact of missingness. The key findings of chapter 4 were:

- Authors of systematic reviews face a number of challenges when trying to identify the number of trial participants with missing outcome data.
- The judgment of missingness relies on how trial authors report on categories of participants and handle them in their analyses.
- We developed a guidance for authors of systematic reviews on how to identify participants with missing outcome data in RCTs.
- The guidance classifies the 19 categories of participants as having either 'definitely missing outcome data' or 'potentially missing outcome data', which has implications on the precision and certainty of results of the meta-analysis.
- For their primary analysis, systematic reviewers must choose between two options: use either 'definitely missing outcome data' or both 'definitely missing outcome data' and 'potentially missing outcome data'. Review authors also need to choose a method for handling missing data in the meta-analysis. To test the robustness of the analysis that follows from these choices, the authors could explore sensitivity analyses using alternatives for identifying participants with missing outcome data and for handling missing outcome data.
- This guidance fits into a larger one that includes first requesting missing data by outcome from trial authors and eventually assessing the associated risk of bias using Grading of Recommendations Assessment, Development and Evaluation.

In **chapter 5**, we assessed risk of bias associated with missing outcome data in systematic reviews by examining how different methods of handling missing data alter statistical significance of pooled effect estimates of dichotomous outcomes and quantifying the change in effect estimate when applying different methods of handling missing outcome data. The key findings of chapter 5 were:

- Almost a quarter of meta-analyses lost significance when using a conservative approach to test their robustness (i.e., applying plausible assumptions to the outcomes of participants with definite missing outcome data).
- When using the same conservative approach, up to a quarter of meta-analyses had a change of at least 18% in their relative effect estimates.

- A substantive percentage of meta-analyses is at serious risk of bias associated with missing outcome data and systematic review authors should rate down the certainty of evidence for risk of bias.

In **chapter 6**, we assessed whether systematic review authors are consistent in their methods of handling missing outcome data across trials included in their meta-analyses, and whether the methods used for handling missing outcome data in their meta-analyses were consistent with the reported methods. The key findings of chapter 6 were:

- Most systematic review authors do not explicitly report their methods to handle missing outcome data.
- Systematic review authors are inconsistent in their methods of handling missing outcome data across their eligible primary trials.
- Of the very few reviews that did explicitly report on their methods, none applied that method consistently across the included trials.

Strengths and limitations

A major strength of this research project is that the chapters of this thesis are interrelated, and each chapter is built on the preceding one. First, the action plans for the five chapters presented in this thesis were based on a previously published protocol (72). In chapters 2 and 3, we explored how authors of systematic reviews and RCTs deal with missing outcome data in terms of reporting, handling, and assessing the risk of bias. These two chapters allowed the development of a list of categories of participants that might have missing outcome data that systematic review authors would frequently encounter. These categories of participants were the basis of the guidance drafted in chapter 4 to help systematic review authors identify participants with missing outcome data. Consequently, we used this guidance in chapter's 5 imputation study to identify participants with missing outcome data and ultimately assess the impact of applying different assumptions on the robustness of the pooled relative effect. In chapter 6, in order to assess whether systematic review authors are consistent in their methods of handling missing outcome data across trials included in their meta-analyses and further assess whether these methods are consistent with the reported methods, we had to refer to the findings of chapters 2, 3 and 4 to retrieve the reported methods. Chapter 2

informed us whether systematic review authors report on a method to handle missing outcome data; chapter 3 informed how trial authors dealt with missing outcome data in their analyses; and chapter 4 provided the guidance upon which we identified participants with missing outcome data.

Another major strength of this research project is that it responded to a need in practice made by the Cochrane Collaboration. This project was supported by the Cochrane Methods Innovation Fund in order to provide Cochrane review authors with specific guidance on how to address participant data in trials included in their reviews. Thus, many experts in the field of trial and systematic review methodology were involved in different phases of the project. Based on the findings of this project, we recommended to update the relevant sections of the Cochrane Handbook and provide detailed suggestions on identifying, reporting, handling, and assessing the risk of bias associated with missing outcome data in the included trials.

Implications for systematic review authors

Better adherence to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement recommendations and the Cochrane Handbook regarding reporting on and handling of missing outcome data in systematic reviews is very crucial (12, 73, 74) . We recommend reviewers to follow the guidance presented in chapter 4 on identifying participants with missing outcome data from RCT reports (71) and apply the selected handling method consistently across all trials included in the meta-analysis. It is very instrumental to report on the previous steps clearly. Moreover, systematic review authors must also not confuse between handling missing outcome data and conducting intention-to-treat analysis. The latter is not a method to handle missing outcome data but to deal effectively with non-compliance in those with available outcome data (34). To assess the risk of bias associated with missing outcome data in a body of evidence, we advise to use specific GRADE guidance on this topic for systematic reviews of both binary and continuous outcomes (56).

Implications for trialists

The first step to addressing missing data in trials is to minimize the extent of - if not avoid - missing outcome data in the first place. For example, we suggest that RCT authors seek to continue following up participants belonging to certain categories (e.g., non-compliant or had adverse events). Many strategies have been suggested to improve retention in RCTs (e.g., monetary incentives) (18-20, 75).

Trial authors must adhere to existing guidance to improve reporting on the proportion, reasons, and mechanisms of missing outcome data, handling missing outcome data in the analysis, and assessing the associated risk of bias in their trials (15, 32, 38, 68). Ideally, trialists would report, in a standardized data file compatible with meta-analysis software, the number of participants randomized, the number of participants with missing outcome data, and the number of events for each outcome and for each arm. We acknowledge that trialists are limited by word count in their journal publications, and that such information may appear in an appendix.

As per the GRADE guidance, we recommend conducting a complete case analysis in the primary analysis (56). However, if investigators have strong hypotheses regarding the direction and magnitude of bias associated with missing outcome data, we recommend that RCT authors apply plausible assumptions about the outcomes of these participants in the primary analysis (56, 76). These assumptions might depend on several factors including but not limited to the question being examined by the RCT, baseline characteristics of the included participants, the nature of the intervention, and the reason for missingness. In terms of assessing the risk of bias associated with missing outcome data, we emphasize performing sensitivity analyses to explore the robustness of the results presented in the main analysis.

Implications for journal editors

The adherence of systematic review authors and trial authors to the above recommendations should be ensured by journal editors. Strict rules and regulations to secure appropriate reporting of missing outcome data would decrease the lack of clarity

on missing outcome data in publications. Another suggestion for journal editors is to promote data sharing through an open science framework. The World Health Organization and the International Committee of Medical Journal Editors have highlighted the importance of sharing RCT data (69). Through open science, trial data would be more available for systematic review authors which allows them to better identify, report on, and handle missing outcome data instead making judgements on the extent of missing outcome data of participants. Even if trialists won't follow up all their trials participant, the availability of trial data might alleviate some of the burden of the missing outcome data of trial participants

Implications for research

Many systematic review authors may not have a solid statistical background. Thus, developing a rule of thumb could guide systematic review authors on how to judge risk of bias associated with missing outcome data at the trial level. Such rule of thumb would account for factors such as (1) percentage of missing outcome data per arm, (2) ratio of missing outcome data to event rate per arm, (3) fragility of statistical significance (i.e., borderline significance), (4) magnitude of the effect estimate, and (5) duration of follow-up.

This research project is based on group-level meta-analyses which provide limited information. Reproducing chapters 3, 4, and 5 using individual participant data would elucidate how trial authors actually follow-up and handle participants belonging to certain categories of missing outcome data (chapter 3) which allows verifying the guidance suggested in chapter 4. In addition, individual participant data would allow testing other imputation methods (e.g., multiple imputations) in order to assess the potential impact of missing outcome data on the results of systematic reviews (chapter 5).

References

1. 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.
2. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009;151(4):264-9, W64.
3. The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions2011 March 2011.
4. Kahale LA, Guyatt GH, Agoritsas T, Briel M, Busse JW, Carrasco-Labra A, et al. A guidance was developed to identify participants with missing outcome data in randomized controlled trials. Journal of Clinical Epidemiology. 2019.
5. 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(11):e49163.
6. 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.
7. Brueton VC, Tierney J, Stenning S, Harding S, Meredith S, Nazareth I, et al. Strategies to improve retention in randomised trials. The Cochrane database of systematic reviews. 2013(12):1.
8. Dziura JD, Post LA, Zhao Q, Fu Z, Peduzzi P. Strategies for dealing with missing data in clinical trials: from design to analysis. Yale J Biol Med. 2013;86(3):343-58.
9. Hollestein L, Carpenter J. Missing data in clinical research: an integrated approach. British Journal of Dermatology. 2017;177(6):1463-5.
10. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al. The prevention and treatment of missing data in clinical trials. N Engl J Med. 2012;367(14):1355-60.
11. Akl EA, Shawwa K, Kahale LA, Agoritsas T, Brignardello-Petersen R, Busse JW, et al. Reporting missing participant data in randomised trials: systematic survey of the methodological literature and a proposed guide. BMJ Open. 2015;5(12):e008431.
12. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux P, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Bmj. 2010;340:c869.
13. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Int J Surg. 2012;10(1):28-55.
14. National Research Council. The Prevention and Treatment of Missing Data in Clinical Trials. Washington (DC): The National Academic Press; 2010.
15. White IR, Carpenter J, Horton NJ. Including all individuals is not enough: lessons for intention-to-treat analysis. Clin Trials. 2012;9(4):396-407.
16. Naudet F, Sakarovitch C, Janiaud P, Cristea I, Fanelli D, Moher D, et al. Data sharing and reanalysis of randomized controlled trials in leading biomedical journals with a full data sharing policy: survey of studies published in The BMJ and PLOS Medicine. bmj. 2018;360:k400.

Appendices

Summary

Missing outcome data of trial participants is a frequent phenomenon in RCTs and may represent a serious potential source of bias if not reported and handled appropriately. The potential effect of bias associated with missing outcome data- attrition bias- is that invalid conclusions about efficacy and safety of studied interventions may be reached and ultimately impact clinical practice. The poor reporting and handling of missing outcome data in RCTs contribute to the inadequate reporting and handling of missing outcome data in systematic reviews. Systematic review authors need specific guidance on how to address missing outcome data of trial participants in their reviews. The overall aim of this thesis is to address this need. Specifically, we aim to:

1. Describe how systematic review authors report on the categories of participants who might have missing outcome data, handle missing outcome data in their primary meta-analyses of dichotomous outcomes, and assess the associated risk of bias;
2. Assess how trial authors report on the categories of participants that might have missing outcome data and their follow-up status, and on the handling of these participants in their main and secondary analyses;
3. Provide guidance for authors of systematic reviews on how to identify and classify participants with missing outcome data in trials;
4. Assess risk of bias associated with missing outcome data in systematic reviews by examining how different methods of handling missing outcome data alter statistical significance of pooled effect estimates of dichotomous outcomes and quantifying the change in effect estimate when applying different methods of handling missing outcome data;
5. Assess whether systematic review authors are consistent in their methods of handling missing outcome data across trials included in their meta-analyses, and whether the methods used for handling missing outcome data in their meta-analyses were consistent with the reported methods.

In **chapter 2**, we assessed how authors of 100 Cochrane and non-Cochrane systematic review authors report and address categories of participants with who might have missing outcome data. The methodological survey showed that most systematic reviews

do not explicitly report sufficient information on categories of trial participants with potential missing outcome data or address missing outcome data in their primary analyses. Only 19 systematic reviews reported plans for handling missing outcome data in their analyses. Although 87 systematic reviews addressed risk of bias associated with missing outcome data at the trial level, only nine reported conducting sensitivity analysis as a way to judge risk of bias associated with missing outcome data at the level of the meta-analysis.

Chapter 3 surveyed all RCT reports included in the 100 systematic reviews included in chapter 2 to describe how RCT authors (1) report on different categories of participants that might have missing outcome data, (2) handle these categories in the analysis, and (3) judge the risk of bias associated with missing outcome data. The survey showed that the majority of trials did not clearly report on whether different categories of participants that might have missing outcome data have been followed-up or not. The median percentage of participants who were explicitly not followed-up was 5.8% (IQR 2.2-14.8%). When one also includes participants with unclear follow-up status, the total value rises to 11.7% (IQR 5.6-23.7%). In addition, most trials did not specify how they handled missing outcome data in their analysis, did not reported on missing outcome data separately for different outcomes, and did not address risk of bias associated with missing outcome data. Very few RCTs described a mechanism of missingness (e.g., missing completely at random, missing not at random), and none took uncertainty into account when imputing outcomes.

Chapter 4 presented guidance for authors of systematic reviews on how to identify participants with missing outcome data in trial reports, especially when trial reporting is not clear. Our approach was based on how trial authors report on categories of participants who might have missing outcome data and how they handle them in their analyses. When trialists explicitly report that participants were followed-up, then systematic review authors need to count them as definitely not having missing outcome data. When trialists explicitly report that participants were not followed-up, then systematic review authors need to count them as definitely having missing outcome data. In most circumstances, trial authors do not explicitly report follow-up status, thus systematic review authors should check how authors handled these categories. If such

participants were excluded from the trial analysis (i.e., excluded from the denominator (and numerator)) or for whom the trialists imputed outcomes, then systematic reviewer authors should count them definitely missing. Very frequently, it is unclear how primary study investigators handled participants with unclear follow-up status, thus, it would be best to count them as potentially missing outcome data. At the meta-analysis level, systematic reviewers must choose between two options: use either 'definitely missing outcome data' or 'total possible missing outcome data' (which is the combination of both definite and potential missing outcome data). In contrast to definitely missing outcome data, using the 'total possible missing outcome data' in the primary analysis will yield less precision in the pooled effect. However, to test the robustness of the analysis that follows from these choices, the authors could explore sensitivity analyses using alternatives for identifying participants with missing outcome data and for handling missing outcome data.

Chapter 5 explores the potential impact of missing outcome data on effect estimates of the 100 meta-analyses included in chapter 2. We examined how the statistical significance of pooled effect estimates is affected by the different methods of handling missing outcome data. When applying plausible assumptions to the outcomes of participants with definite missing outcome data, up to a quarter of meta-analyses lost statistical significance. When applying implausible but commonly assumptions, the percentage of systematic reviews that lost significance was as high as 60% with the worst-case scenario.

We also explored the magnitude of change in the effect estimate when applying different assumptions to the outcomes of participants with missing outcome data. When applying plausible assumptions to the outcomes of participants with definite missing outcome data, the median change in relative effect estimate was as high as 7.0% (IQR 2.7%-18.2%); when applying implausible but commonly used assumptions, the median change was as large as 30.4% (IQR 10.5%- 77.5%).

Chapter 6 compared how authors of the 100 systematic reviews included in chapter 2 reported to handle missing outcome data of trial participants to what they actually did. We highlighted three major limitations in how systematic reviews authors handle

missing outcome data in their meta-analysis: (1) lack of transparency in reporting the method of dealing with missing outcome data; (2) high inconsistency in the method used in different RCTs within the same meta-analysis; and (3) and high inconsistency between the method used and the method reported.

Finally, in **Chapter 7**, we discussed the main findings of each chapter, discussed the strengths and limitations, and provided implications for practice for systematic review authors, for trialists, for journal editors, and for research. Then main implications are as follows:

- Systematic review authors should better adhere to the PRISMA statement and Cochrane's handbook regarding reporting on, and handling of missing outcome data;
- Systematic review authors need to refer to the guidance presented in chapter 4 on identifying participants with missing outcome data from RCT reports and apply the handling method consistently across all trials included in the meta-analysis;
- Systematic review authors need to refer to the GRADE guidance to assess the risk of bias associated with missing outcome data in a body of evidence;
- Systematic review authors must report clearly and explicitly on the previous steps;
- Trial authors should minimize the extent of - if not avoid missing outcome data in the first place at the level of the trial design;
- Trial authors must adhere to existing guidance to improve reporting on the proportion, reasons, and mechanisms of missing outcome data, handling missing outcome data in the analysis, and assessing the associated risk of bias;
- Trial authors need to report, in a standardized data file compatible with meta-analysis software, the number of participants randomized, the number of participants with missing outcome data, and the number of events for each outcome and for each arm;
- Journal editors should ensure the adherence of systematic review authors and trial authors to the above recommendations;
- journal editors need to promote data sharing through an open science framework;

- There is a need to develop a rule of thumb that could guide systematic review authors on how to judge risk of bias associated with missing outcome data at the trial level;
- developing a rule of thumb could guide systematic review authors on how to judge risk of bias associated with missing outcome data at the trial level;
- There is a need to reproduce the five chapters using systematic reviews of individual participant data.

Nederlandse samenvatting

Ontbrekende resultaten met betrekking tot de in een onderzoek bestudeerde uitkomsten bij deelnemers aan een *randomised controlled trial* (RCT) komen vaak voor. Dergelijke ontbrekende resultaten kunnen een belangrijke bron van vertekening vormen, als ze niet op de juiste manier worden gerapporteerd en behandeld in de analyse. Een mogelijk gevolg van vertekening door ontbrekende resultaten - *attrition bias* - is dat ongeldige conclusies over de werkzaamheid en veiligheid van de bestudeerde interventies worden getrokken, die op hun beurt de klinische praktijk op een verkeerde wijze kunnen beïnvloeden. De gebrekkige rapportage en verwerking van ontbrekende resultaten in RCT's dragen daarnaast bij aan ontoereikende rapportage en verwerking van ontbrekende resultaten in systematische literatuuroverzichten (*systematic reviews* – systematische reviews). Een specifieke leidraad over hoe in systematische reviews om te gaan met deze ontbrekende resultaten zijn echter nog niet voorhanden. Het doel van dit proefschrift was in deze behoefte te voorzien, waarbij wij ons gericht hebben op dichotome uitkomsten (uitkomsten die wel of niet aanwezig zijn). Onze specifieke doelstellingen waren de volgende:

1. Te inventariseren hoe auteurs van systematische reviews rapporteren over verschillende typen deelnemers met mogelijk ontbrekende resultaten, hoe zij ontbrekende resultaten verwerken in hun primaire meta-analyses en hoe zij de bijbehorende kans op vertekening door ontbrekende resultaten inschatten;
2. Te inventariseren hoe onderzoekers die RCT's uitvoeren, rapporteren over verschillende typen onderzoekdeelnemers met mogelijk ontbrekende resultaten en hun follow-upstatus, en hoe zij in hun analyses omgegaan zijn met dergelijke deelnemers;
3. Een leidraad te ontwikkelen voor auteurs van systematische reviews met betrekking tot het identificeren en classificeren van deelnemers met ontbrekende resultaten in de in hun review opgenomen onderzoeken;
4. De kans op vertekening te analyseren door a) te onderzoeken hoe verschillende methoden voor het omgaan met ontbrekende resultaten de statistische significantie van gepoolde effectschattingen van dichotome uitkomsten kunnen veranderen en b) de verandering in effectschatting na toepassing van verschillende imputatiemethoden te kwantificeren;

5. Te onderzoeken of auteurs van systematische reviews a) consistent zijn met betrekking tot de methoden voor het verwerken van ontbrekende resultaten in de in hun meta-analyses opgenomen onderzoeken en b) of de methoden die zij uiteindelijk gebruiken voor het verwerken van ontbrekende resultaten, overeenkomen met hetgeen ze beschreven in hun methodeparagraaf.

In **hoofdstuk 2** onderzochten wij hoe auteurs van 50 Cochrane- en 50 niet-Cochrane reviews rapporteerden over en in hun primaire meta-analyses omgingen met verschillende typen deelnemers met mogelijk ontbrekende resultaten. In het merendeel van de systematische reviews werd hieraan onvoldoende aandacht besteed. Slechts in 19 systematische reviews werden de methoden beschreven voor het verwerken van ontbrekende resultaten in de meta-analyses. Hoewel in 87 systematische reviews de kans op vertekening door ontbrekende resultaten in de in de review opgenomen RCT's onderzocht werd, werden slechts in negen reviews zogenoemde sensitiviteitsanalyses uitgevoerd om de kans op vertekening op het niveau van de meta-analyse te beoordelen.

In **hoofdstuk 3** bekeken wij de afzonderlijke publicaties van alle RCT's die in de 100 systematische reviews van hoofdstuk 2 opgenomen waren. Wij inventariseerden, (1) hoe onderzoekers rapporteerden over deelnemers met mogelijk ontbrekende resultaten, (2) hoe zij met de verschillende typen van dergelijke deelnemers omgegaan zijn in hun analyses en (3) hoe zij de kans op vertekening door ontbrekende resultaten hebben ingeschat. In het merendeel van de onderzoeken werd niet duidelijk gerapporteerd of deelnemers met mogelijk ontbrekende resultaten al dan niet zijn opgevolgd. Het mediane percentage deelnemers van wie duidelijk was dat zij niet werden opgevolgd, bedroeg 5,8% (interkwartielafstand 2,2-14,8%). Dit percentage steeg tot 11,7% (interkwartielafstand 5,6-23,7%) als ook deelnemers met een onduidelijke follow-upstatus werden meegenomen in deze analyse. In de meeste onderzoeken werd niet beschreven, hoe de onderzoekers met ontbrekende resultaten in hun analyse omgingen, werden ontbrekende resultaten niet afzonderlijk voor de verschillende bestudeerde uitkomsten gerapporteerd en werd geen inschatting gemaakt van de kans op vertekening door ontbrekende resultaten. Zeer weinig RCT's beschreven de mogelijke aard van ontbrekende resultaten (bijv. *missing completely at random*, *missing not at random*), en

geen ervan hield rekening met mogelijke onzekerheid bij het imputeren van ontbrekende resultaten.

In **hoofdstuk 4** ontwikkelden wij een leidraad voor auteurs van systematische reviews voor het identificeren van deelnemers met ontbrekende resultaten in de in hun review opgenomen onderzoeken, met name in situaties waarin hierover in de afzonderlijke onderzoeken niet duidelijk gerapporteerd wordt. Wij zijn hierbij uitgegaan van de wijze waarop de primaire onderzoekers rapporteerden over verschillende typen deelnemers met mogelijk ontbrekende resultaten en hoe zij daarmee omgingen in hun analyses. Wanneer onderzoekers expliciet vermelden dat de follow-up van alle deelnemers volledig was, dan kunnen systematische reviewauteurs deze beschouwen als 'zeker niet ontbrekend'. Wanneer onderzoekers expliciet rapporteren dat niet bij alle deelnemers de follow-up volledig was, dan kunnen systematische reviewauteurs deze als 'zeker ontbrekend' beschouwen. In de meeste gevallen wordt de follow-upstatus van de deelnemers echter niet expliciet gerapporteerd. Systematische reviewauteurs zouden dan moeten nagaan op welke wijze de onderzoekers zijn omgegaan met deze deelnemers. Als dergelijke deelnemers werden uitgesloten van de analyse (d.w.z. uitgesloten van de noemer (en teller)) of als de onderzoekers voor deze deelnemers resultaten imputeerden, dan zouden systematische reviewauteurs deze eveneens als 'zeker ontbrekend' moeten beschouwen. Heel vaak zal echter niet duidelijk zijn hoe de onderzoekers omgingen met deelnemers met een onduidelijke follow-upstatus. In dergelijke gevallen kunnen deze deelnemers het beste als 'mogelijk ontbrekend' geduid worden. Voor hun meta-analyse zouden systematische reviewauteurs dan moeten kiezen tussen twee opties: verwerk alleen de 'zeker ontbrekende' resultaten of neem hiervoor zowel de zeker ontbrekende resultaten als de mogelijk ontbrekende ('totaal mogelijk ontbrekend'). Deze laatste optie zal minder precieze gepoolde effectschattingen opleveren (bredere betrouwbaarheidsintervallen). Om de invloed te onderzoeken van de keuze voor de analyse kunnen de systematische reviewauteurs sensitiviteitsanalyses uitvoeren waarin beide methoden worden vergeleken.

In **hoofdstuk 5** onderzochten wij de mogelijke invloed van ontbrekende resultaten op de gepoolde effectschattingen van de 100 in hoofdstuk 2 opgenomen meta-analyses. Wij gingen na hoe de statistische significantie beïnvloed werd door de methode voor het

omgaan met ontbrekende resultaten. Bij het toepassen van plausibele aannames over wat de uitkomst zou kunnen zijn bij deelnemers met ontbrekende resultaten, werd bij bijna een kwart van de meta-analyses het resultaat niet significant. Bij gebruik van vaak toegepaste, meer extreme, maar niet erg plausibele aannames, zoals het *worst-case scenario*, werd het resultaat in bijna 60% van de meta-analyses niet significant.
We onderzochten ook in welke mate de grootte van de gepoolde effectschatting veranderde door verschillende aannames met betrekking tot de uitkomsten van deelnemers met ontbrekende resultaten. Bij het toepassen van de eerder genoemde plausibele aannames werd de schatting van het relatieve risico mediaan 7,0% hoger (interkwartielafstand 2,7% tot 18,2%). Bij het toepassen van de gangbare, meer extreme, maar niet erg plausibele aannames bedroeg deze mediane verandering 30,4% (interkwartielafstand 10,5% tot 77,5%).

In **hoofdstuk 6** vergeleken wij wat de auteurs van de 100 in hoofdstuk 2 opgenomen systematisch reviews in hun methodeparagraaf vermeldden over het omgaan met deelnemers met ontbrekende resultaten met wat ze daadwerkelijk deden in hun meta-analyses. We stelden hierbij de volgende drie tekortkomingen vast: (1) onduidelijke of geen vermelding van de methode voor het omgaan met ontbrekende resultaten; (2) aanzienlijke inconsistentie met betrekking tot de methode die werd gebruikt voor de verschillende onderzoeken binnen dezelfde meta-analyse; en (3) aanzienlijke inconsistentie tussen de geplande methode en de daadwerkelijk toegepaste methode .

In **hoofdstuk 7** ten slotte vatten we de belangrijkste bevindingen van elk hoofdstuk samen, bespraken wij de sterke punten en beperkingen van ons onderzoek en gaven wij aanbevelingen voor de praktijk voor diverse doelgroepen. De belangrijkste aanbevelingen zijn de volgende:

Auteurs van systematische reviews

- dienen zich beter te houden aan de PRISMA rapportage-eisen en de methoden voor het omgaan met ontbrekende resultaten, zoals beschreven in het *Cochrane Handbook for Systematic Reviews of Interventions*;
- dienen de in hoofdstuk 4 ontwikkelde leidraad voor het identificeren van deelnemers met ontbrekende resultaten toe te passen en de manier van omgaan met

onderzoekdeelnemers met ontbrekende resultaten consistent te gebruiken in alle in de meta-analyses opgenomen onderzoeken;
- dienen de GRADE-systematiek te gebruiken om – per uitkomst – de kans op vertekening ten gevolge van ontbrekende resultaten in te schatten;
- dienen de voorgaande stappen duidelijk en expliciet te rapporteren.

Onderzoekers die zelf gerandomiseerde studies uitvoeren
- dienen in hun onderzoekontwerp maatregelen op te nemen om ontbrekende resultaten zoveel mogelijk te beperken (d.w.z. een volledige follow-up te bewerkstelligen), zo niet geheel te voorkomen;
- dienen zich te houden aan de bestaande rapportage-eisen over de omvang van, redenen voor en onderliggende mechanismen betreffende ontbrekende resultaten, en dienen ontbrekende resultaten op een methodologisch verantwoorde wijze in hun analyse te verwerken en de bijbehorende kans op vertekening in te schatten;
- dienen voor iedere uitkomst en voor iedere onderzoeksarm het aantal gerandomiseerde deelnemers te vermelden, het aantal deelnemers met ontbrekende resultaten en het aantal deelnemers bij wie de (dichotome) uitkomst van interesse aanwezig was.

Redacteuren van medisch-wetenschappelijke tijdschriften
- moeten ervoor zorgdragen dat systematische reviewauteurs en primaire onderzoekers zich houden aan bovenstaande aanbevelingen;
- moeten het openbaar maken en delen van onderzoekgegevens bevorderen.

Aanbevelingen voor verder onderzoek:
- Er moet een vuistregel ontwikkeld worden die systematische reviewauteurs helpt bij het inschatten van de kans op vertekening ten gevolge van ontbrekende resultaten op het niveau van ieder onderzoek.
- De onderzoeken die wij in de hoofdstukken 2 t/m 6 uitvoerden, betroffen systematische reviews van geaggregeerde resultaten. Om de resultaten van onze onderzoeken te bevestigen, zouden deze herhaald moeten worden met systematische reviews van individuele deelnemersgegevens (IPD meta-analyses).

Acknowledgment

The thesis work would have been impossible without enormous physical, emotional, and intellectual support and help from many individuals along the way. I would like to thank them and express my deepest appreciation.

I would like to express my sincere gratitude to Dr. Rob Scholten, for his constant support, advice, and guidance throughout this thesis project. Without him, this PhD would have not been possible. It has been an honor to work with him.

I would love to express my most special thanks to Dr. Elie Akl for his incredible support, guidance, supervision, and mentorship along the way. He has been a role model for my life in so many aspects. I deeply appreciate the opportunity to work with him and all the great things I have learnt from him in and out of work.

I would also like to thank Dr. Gordon Guyatt for sharing his immense experience in health research methodology and providing critical feedback and Dr. Lotty Hooft for her bright ideas that always stimulated my critical thinking.

I have to thank the team of friends and colleagues who worked with me along the way for all of their extremely hard work, effort, dedication, and the time they invested in this project.

My family has provided enormous amount of support, patience, and motivation throughout this journey till this day, for them I am very indebted.

Above all, I thank God for everything.

Curriculum Vitae

Lara A Kahaleh was born on the 30th of November 1989 in Beirut, Lebanon. She completed her Bachelor of Science in Nursing from the University of Balamand in 2011. She worked as a registered nurse at the Neonatal Intensive Care Unit in 2012 then shifted to the field of health research in 2013 at the Clinical Research Institute at the American University of Beirut, Beirut, Lebanon. The same year, Lara received a certificate from the Scholars in Health Research Program at the American University of Beirut in collaboration with Harvard Medical School and Harvard School of Public Health funded by National Institutes of Health. In 2015, she pursued her Master's in Epidemiology at the American University of Beirut. Her thesis topic was about assessing risk of bias associated with missing dichotomous outcome data in meta-analyses: application in five Cochrane systematic reviews, which is a pilot study of chapter 5 of this PhD thesis. In 2016, Lara started her PhD studies at the Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, The Netherlands, under the supervision of Professors Dr. Rob Scholten from Utrecht University (The Netherlands), Dr. Elie Akl from the American University of Beirut (Lebanon), and Dr. Lotty Hooft from Utrecht University (The Netherlands), and under the mentorship of Professor Dr. Gordon Guyatt from McMaster University (Canada).

Lara currently coordinates the AUB GRADE centre at the American University of Beirut. She has been involved in the conduct of more than 40 systematic reviews and methodological studies since 2012. She is currently coordinating the development of guidelines for the American Society of Haematology and the American College of Rheumatology. She coordinated the adaptation of clinical guidelines as well for the Brazilian and Tunisian Ministries of Health. Her research interests are in systematic review methodology, missing outcome data, and clinical practice guidelines. She is a member of several Cochrane groups, the GRADE working group, and the Lebanese Epidemiological association.

List of publications

Publications in this thesis

1. **Kahale LA**, Guyatt GH, Agoritsas T, Briel M, Busse JW, Carrasco-Labra A, Khamis AM, Zhang Y, Hooft L, Scholten RJ, Akl EA, A guidance was developed to identify participants with missing outcome data in randomized controlled trials. **Journal of Clinical Epidemiology** (2019), doi: https://doi.org/10.1016/j.jclinepi.2019.07.003.

2. **Kahale LA**, Diab B, Khamis AM, Chang Y, Lopes LC, Agarwal A, Li L, Mustafa R, Koujanian S, Waziry R, Busse JW, Dakik A, Guyatt G, and Akl EA. Potentially missing data was considerably more frequent than definitely missing data in randomized controlled trials: A methodological survey. **Journal of Clinical Epidemiology**. 2018 Oct 6.

3. **Kahale LA**, Diab B, Brignardello-Petersen R, Agarwal A, Mustafa RA, Kwong J, Neumann I, Li L, Lopes LC, Briel M, Busse JW, Iorio A, Vandvik PO, Alexander PE, Guyatt G, and Akl EA. Systematic reviews do not adequately report or address missing outcome data in their analyses: a methodological survey. **Journal of Clinical Epidemiology**. 2018 Jul 1;99:14-23.

Submitted manuscripts

4. **Kahale LA,** Khamis AM, Diab B, Chang Y, Lopes LC, Agarwal A, Li L, Mustafa R, Koujanian S, Waziry R, Busse JW, Dakik A, Schünemann H, Hooft L, Guyatt G, Scholten RJPM, and Akl EA. Potential impact of missing outcome data on treatment effects in systematic reviews: an imputation study. **BMJ** (submitted)**.**

5. **Kahale LA**, Khamis AM, Diab B, Chang Y, Lopes LC, Agarwal A, Li L, Mustafa R, Koujanian S, Waziry R, Busse JW, Dakik A, Schünemann H, Hooft L, Guyatt G, Scholten RJPM, and Akl EA. Meta-analyses proved inconsistent in how missing data were handled across their included primary trials: a methodological survey. **International Journal of Epidemiology** (submitted).

6. Surdyk S, Itani M, Al-Lobaidy M, **Kahale LA**, Farha A, Dewachi O, Akl EA, and Habib RR. Weaponized uranium and adverse health outcomes in Iraq: A systematic review Environmental Health Perspectives. **Environmental Health Perspectives** (submitted).

7. Abboud J, Dempster M, Rahman AA, **Kahale LA,** Adair P. Prevention of health care associated venous thromboembolism through implementing VTE prevention clinical practice guidelines in hospitalized medical patients: A systematic review and meta-analysis. **Implementation Science** (submitted).

Additional publications

8. Atkinson KM, Wilson K, Murphy MS, El-Halabi S, **Kahale LA**, Laflamme LL, El-Khatib Z. Effectiveness of digital technologies at improving vaccine uptake and series completion–A systematic review and meta-analysis of randomized controlled trials. **Vaccine**. 2019 Apr 26.

9. Khamis AM, **Kahale LA**, Pardo-Hernandez H, Schünemann HJ, Akl EA. Methods of conduct and reporting of living systematic reviews: a protocol for a living methodological survey. **F1000Research** 2019, 8:221 (https://doi.org/10.12688/f1000research.18005.1).

10. Matar CF, **Kahale LA**, Hakoum MB, Tsolakian IG, Etxeandia-Ikobaltzeta I, Yosuico VED, Terrenato I, Sperati F, Barba M, Schünemann H, Akl EA. Anticoagulation for perioperative thromboprophylaxis in people with cancer. **Cochrane Database of Systematic Reviews** 2018, Issue 7. Art. No.: CD009447. DOI: 10.1002/14651858.CD009447.pub3 *(Highlighted in Cochrane Clinical Answers 'How does low-molecular-weight heparin (LMWH) compare with unfractionated heparin (UFH) or fondaparinux for perioperative thromboprophylaxis in people with cancer?' Simone Mocellin (MD) (on behalf of Cochrane Clinical Answers Editors). Cochrane Clinical Answers 2018. DOI: 10.1002/cca.2256)*

11. **Kahale LA**, Hakoum MB, Tsolakian IG, Alturki F, Matar CF, Terrenato I, Sperati F, Barba M, Yosuico VED, Schünemann H, Akl EA. Anticoagulation for the long-term treatment of venous thromboembolism in people with cancer. **Cochrane Database of Systematic Reviews** 2018, Issue 6. Art. No.: CD006650. DOI: 10.1002/14651858.CD006650.pub5 *(Highlighted in Al-Samkari H. Optimal anticoagulant treatment of cancer-associated venous thromboembolism remains unclear. BMJ evidence-based medicine. 2018 Dec 5:bmjebm-2018 and in Cochrane Clinical Answers 'How do low-molecular-weight heparin (LMWH), vitamin K agonists (VKAs), and direct oral anticoagulants (DOACs) compare for treatment of venous thromboembolism (VTE) in people with cancer?' Simone Mocellin (MD) (on behalf of Cochrane Clinical Answers Editors). Cochrane Clinical Answers 2018. DOI: 10.1002/cca.2255).*

12. **Kahale LA**, Tsolakian IG, Hakoum MB, Matar CF, Barba M, Yosuico VED, Terrenato I, Sperati F, Schünemann H, Akl EA. Anticoagulation for people with cancer and central venous catheters. **Cochrane Database of Systematic Reviews** 2018, Issue 6. Art. No.: CD006468. DOI: 10.1002/14651858.CD006468.pub6 *(Highlighted in Cochrane Clinical Answers 'What are the benefits and harms of anticoagulation for people with cancer and central venous catheters?' Maitreyee Rai (MBBS) (on behalf of Cochrane Clinical Answers Editors). Cochrane Clinical Answers 2018. DOI: 10.1002/cca.2228).*

13. Hakoum MB, **Kahale LA**, Tsolakian IG, Matar CF, Yosuico VED, Terrenato I, Sperati F, Barba M, Schünemann H, Akl EA. Anticoagulation for the initial treatment of venous thromboembolism in people with cancer. **Cochrane Database of Systematic Reviews** 2018, Issue 1. Art. No.: CD006649. DOI: 10.1002/14651858.CD006649.pub7.

14. **Kahale LA**, Hakoum MB, Tsolakian IG, Matar CF, Barba M, Yosuico VE, Terrenato I, Sperati F, Schünemann H, Akl EA. Oral anticoagulation in people with cancer who have no therapeutic or prophylactic indication for anticoagulation. **Cochrane Database of Systematic Reviews** 2017 Dec; 12:CD006466.

15. Abdel Rahman A, Jomaa L, **Kahale LA**, Adair P, Pine C. Effectiveness of behavioral interventions to reduce the intake of sugar-sweetened beverages in children and adolescents: a systematic review and meta-analysis. **Nutrition reviews**. 2017 Dec 21.

16. Fadlallah R, Al-Khaled L, Brax H, Nasser M, Rajabbik M, Nass H, **Kahale LA**, Akl EA. Extent of physician pharmaceutical industry interactions in low- and middle-income countries: a systematic review. **European Journal of Public Health**. 2017 Nov 20. Doi: 10.1093/eurpub/ckx204.

17. Simmonds M, Salanti G, **et al**. Living systematic reviews: 3. Statistical methods for updating meta-analyses. **Journal of Clinical Epidemiology**. 2017 Nov; 91:38-46. Doi: 10.1016/j.jclinepi.2017.08.008.

18. Thomas J, Noel-Storr A, **et al**. 'Living systematic reviews: 2. Combining human and machine effort.' **Journal of Clinical Epidemiology** 2017 Nov; 91:31-37. Doi: 10.1016/j.jclinepi.2017.08.011.

19. Elliott JH, Synnot A, **et al**. Living systematic review: 1. Introduction—the why, what, when, and how. **Journal of Clinical Epidemiology**. 2017 Nov; 91:23-30. Doi: 10.1016/j.jclinepi.2017.08.010.

20. Akl EA, Meerpohl JJ, Elliott J, **Kahale LA**, Schünemann HJ on behalf of the Living Systematic Review Network. Living systematic reviews: 4. Living guideline recommendations. **Journal of Clinical Epidemiology**. September 2017. DOI: 10.1016/j.jclinepi.2017.08.009

21. Akl EA, **Kahale LA**, Hakoum MB, Matar CF, Sperati F, Barba M, Yosuico VED, Terrenato I, Synnot A, Schünemann H. Parenteral anticoagulation in ambulatory patients with cancer. **Cochrane Database of Systematic Reviews** 2017, Issue 9. Art. No.:CD006652. DOI: 10.1002/14651858.CD006652.pub5.

22. Zhang Y, Alyass A, Vanniyasingam T, Sadeghirad B, Flórez ID, Pichika SC, Kennedy SA, Abdulkarimova U, Zhang Y, Iljon T, Morgano GP, Colunga Lozano LE, Aloweni FAB, Lopes LC, Yepes-Nuñez JJ, Fei Y, Wang L, **Kahale LA**, Meyre D, Akl EA, Thabane L, Guyatt GH. A systematic survey of the methods literature on the reporting quality and optimal methods of handling participants with missing outcome data for continuous outcomes in randomized controlled trials. **Journal of Clinical Epidemiology**. 2017 Aug 1;88:67-80.

23. Zhang Y, Flórez ID, Colunga Lozano LE, Bakar Aloweni FA, Kennedy SA, Li A, Craigie S, Zhang S, Agarwal A, Lopes LC, Devji T, Wiercioch W, Riva JJ, Wang M, Jin X, Fei Y, Alexander P, Morgano GP, Zhang Y, Carrasco-Labra A, **Kahale LA**, Akl EA, Schünemann HJ, Thabane L, Guyatt G. A systematic survey on reporting and methods

for handling missing participant data for continuous outcomes in randomized controlled trials**. Journal of Clinical Epidemiology**. 2017 Aug 1; 88:57-66.

24. Guyatt GH, Ebrahim S, Alonso-Coello P, Johnston BC, Mathioudakis AG, Briel M, Mustafa RA, Sun X, Walter SD, Heels-Ansdell D, Neumann I, **Kahale LA**, Iorio A, Meerpohl J, Schünemann HJ, Akl EA. GRADE guidelines 17: Assessing the Risk of Bias Associated with Missing Participant Outcome Data in a body of evidence. **Journal of Clinical Epidemiology**. 2017 May 18.

25. Brax H, Fadlallah R, Al-Khaled L, **Kahale LA**, Nas H, El-Jardali F, et al. (2017) Association between physicians' interaction with pharmaceutical companies and their clinical practices: A systematic review and metaanalysis. **PLoS ONE** 12(4): e0175493. https://doi.org/ 10.1371/journal.pone.0175493.

26. Ballout RA, Foster JP, **Kahale LA**, Badr L. Body positioning for spontaneously breathing preterm infants with apnoea. **Cochrane Database of Systematic Reviews** 2017, Issue 1. Art. No.: CD004951. DOI: 10.1002/14651858.CD004951.pub3.

27. Fadlallah R, Nas H, Naamani D, El-Jardali F, Hammoura I, Al-Khaled L, Brax H, **Kahale LA**, Akl EA. Knowledge, Beliefs and Attitudes of Patients and the General Public towards the Interactions of Physicians with the Pharmaceutical and the Device Industry: A Systematic Review. **PLoS ONE** 2016, 11(8): e0160540. doi: 10.1371/journal.pone.0160540.

28. Akl EA, **Kahale LA**, Ebrahim S, Alonso–Coello P, Schünemann HJ, Guyatt GH. Three challenges described for identifying participants with missing data in trials reports, and potential solutions suggested to systematic reviewers. **Journal of Clinical Epidemiology** 2016, doi: 10.1016/j.jclinepi.2016.02.022.

29. Lotfi T, Morsi RZ, Rajabbik MH, Alkhaled L**, Kahale L**, Nass H, Brax H, Fadlallah R, Akl EA. Knowledge, beliefs and attitudes of physicians in low and middle–income countries regarding interacting with pharmaceutical companies: a systematic review. **BMC Health Services Research** 2016, 16(57), doi 10.1186/s12913–016–1299–4.

30. Akl EA, Shawwa K, **Kahale LA**, Agoritsas T, Brignardello–Petersen R, Busse JW, Carrasco–Labra A, Ebrahim S, Johnston BC, Neumann I, Sola I, Sun X, Vandvik P, Zhang Y, Alonso–Coello P, Guyatt GH. Reporting missing participant data in randomised trials: systematic survey of the methodological literature and a proposed guide. **BMJ Open** 2015;5: e008431 doi:10.1136/bmjopen–2015–008431.

31. Lotfi T, Morsi ZR, Zmeter N, Godah MW, Alkhaled L, **Kahale LA**, Nass H, Brax H, Fadlallah R, Akl EA. Validity of tools used for surveying physicians about their interactions with pharmaceutical company: a systematic review. **BMC research notes** 2015, 8(720), doi 10.1186/s13104–015–1709–4.

32. Badr LK, Abdallah B, **Kahale L.** A Meta–Analysis of Preterm Infant Massage: An Ancient Practice with Contemporary Applications. **The American Journal of Maternal Child Nursing** 2015, 40(6), 344-358.

33. Akl EA, **Kahale LA**, Agoritsas T, Brignardello–Petersen R, Busse J, Carrasco–Labra A, Ebrahim S, Johnston BC, Neumann I, Sola I, Sun X, Vandvik P, Zhang Y, Alonso–Coello P, Guyatt G. Handling trial participants with missing outcome data when conducting a meta–analysis: a systematic survey of proposed approaches. **Systematic Reviews** 2015, 4:98 doi: 10.1186/s13643–015–0083–6.

34. Akl EA, **Kahale LA**, Schünemann H. Association Between Perioperative Low–Molecular–Weight Heparin vs Unfractionated Heparin and Clinical Outcomes in Patients with Cancer Undergoing Surgery. **JAMA** 2015, 313(13), 1364–1365.

35. Akl EA, **Kahale LA**, Ballout RA, Barba M, Yosuico VED, van Doormaal FF, Middeldorp S, Bryant A, Schünemann H. Parenteral anticoagulation in ambulatory patients with cancer. **Cochrane Database of Systematic Reviews** 2014, Issue 12. Art. No.: CD006652. DOI: 10.1002/14651858.CD006652.pub4.

36. Akl EA, **Kahale LA**, Agarwal A, Al–Matari N, Ebrahim S, Alexander PE, Briel M, Petersen RB, Busse J, Diab B, Iorio A, Kwong J, Li L, Lopes LC, Mustafa R, Neumann I, Tikkinen KAO, Vandvik PO, Zhang Y, Pablo Alonso– Coello P, Guyatt G. Impact of missing participant data for dichotomous outcomes on pooled effect estimates in systematic reviews: a protocol for a methodological study. **Systematic Reviews** 2014, 3:137. Doi: 10.1186/2046–4053–3–137.

37. Akl EA, Ramly EP, **Kahale LA**, Yosuico VED, Barba M, Sperati F, Cook D, Schünemann H. Anticoagulation for people with cancer and central venous catheters. **Cochrane Database of Systematic Reviews** 2014, Issue 10. Art. No.: CD006468. DOI: 10.1002/14651858.CD006468.pub5.

38. Akl EA, **Kahale LA**, Terrenato I, Neumann I, Yosuico VED, Barba M, Sperati F, Schünemann H. Oral anticoagulation in patients with cancer who have no therapeutic or prophylactic indication for anticoagulation. **Cochrane Database of Systematic Reviews** 2014, Issue 6. Art. No.: CD006466. DOI:10.1002/14651858.CD006466

39. Akl EA, **Kahale LA**, Neumann I, Barba M, Sperati F, Terrenato I, Muti P, Schünemann H. Anticoagulation for the initial treatment of venous thromboembolism in patients with cancer. **Cochrane Database of Systematic Reviews** 2014, Issue 6. Art. No.: CD006649. DOI: 10.1002/14651858.CD006649.pub6.

40. Akl EA, **Kahale LA**, Sperati F, Neumann I, Labedi N, Terrenato I, Barba M, Sempos EV, Muti P, Cook D, Schünemann H. Low molecular weight heparin versus unfractionated heparin for perioperative thromboprophylaxis in patients with cancer. **Cochrane Database of Systematic Reviews** 2014, Issue 6. Art. No.: CD009447. DOI: 10.1002/14651858.CD009447.pub2.

41. Akl EA, **Kahale LA**, Barba M, Neumann I, Labedi N, Terrenato I, Sperati F, Muti P, Schünemann H. Anticoagulation for the long–term treatment of venous thromboembolism in patients with cancer. **Cochrane Database of Systematic**

Reviews 2014, Issue 7. Art. No.: CD006650. DOI: 10.1002/14651858. CD006650.pub4.

42. Alkhaled L, **Kahale LA**, Nass H, Brax H, Fadlallah R, Badr K, Akl EA. Legislative, educational, policy and other interventions targeting physicians' interaction with pharmaceutical companies: a systematic review. **BMJ Open** 2014; 4: e004880. doi:10.1136/bmjopen–2014– 004880.