RESEARCH

Evolution of poor reporting and inadequate methods over time in 20 920 randomised controlled trials included in Cochrane reviews: research on research study

Agnes Dechartres *associate professor*^{1,4}, Ludovic Trinquart *researcher*^{1,4}, Ignacio Atal *data scientist*^{1,4}, David Moher *senior scientist*⁵, Kay Dickersin *professor*⁶, Isabelle Boutron *professor*^{1,4}, Elodie Perrodeau *statistician*^{1,3}, Douglas G Altman *professor*⁷, Philippe Ravaud *professor*^{1,4,8}

¹INSERM, U1153, Paris, France; ²Faculté de Médecine, Université Paris Descartes, Sorbonne Paris Cité, Paris, France; ³Centre d'Épidémiologie Clinique, Hôpital Hôtel Dieu, AP-HP (Assistance Publique des Hôpitaux de Paris), Paris, France; ⁴Cochrane France, Paris, France; ⁵Clinical Epidemiology Program, Ottawa Hospital Research Institute, School of Epidemiology, Public Health and Preventive Medicine, Canadian EQUATOR Centre, University of Ottawa, Ottawa, ON, Canada; ⁶Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA; ⁷Centre for Statistics in Medicine, Oxford, UK; ⁸Columbia University, Mailman School of Public Health, Department of Epidemiology, New York, NY, USA

Abstract

Objective To examine how poor reporting and inadequate methods for key methodological features in randomised controlled trials (RCTs) have changed over the past three decades.

Design Mapping of trials included in Cochrane reviews.

Data sources Data from RCTs included in all Cochrane reviews published between March 2011 and September 2014 reporting an evaluation of the Cochrane risk of bias items: sequence generation, allocation concealment, blinding, and incomplete outcome data.

Data extraction For each RCT, we extracted consensus on risk of bias made by the review authors and identified the primary reference to extract publication year and journal. We matched journal names with *Journal Citation Reports* to get 2014 impact factors.

Main outcomes measures We considered the proportions of trials rated by review authors at unclear and high risk of bias as surrogates for poor reporting and inadequate methods, respectively.

Results We analysed 20 920 RCTs (from 2001 reviews) published in 3136 journals. The proportion of trials with unclear risk of bias was 48.7% for sequence generation and 57.5% for allocation concealment; the proportion of those with high risk of bias was 4.0% and 7.2%, respectively. For blinding and incomplete outcome data, 30.6% and 24.7% of trials were at unclear risk and 33.1% and 17.1% were at high risk, respectively. Higher journal impact factor was associated with a lower proportion of trials at unclear or high risk of bias. The proportion of trials at unclear risk of bias decreased over time, especially for sequence generation, which fell from 69.1% in 1986-1990 to 31.2% in 2011-14 and for allocation concealment (70.1% to 44.6%). After excluding trials at unclear risk of bias, use of inadequate methods also decreased

over time: from 14.8% to 4.6% for sequence generation and from 32.7% to 11.6% for allocation concealment.

Conclusions Poor reporting and inadequate methods have decreased over time, especially for sequence generation and allocation concealment. But more could be done, especially in lower impact factor journals.

Introduction

The public has the right to expect that information about the efficacy and safety of health interventions is complete, transparent, and reliable and that research investments are not wasted.¹⁻⁷ Randomised controlled trials (RCTs) are the reference standard for assessing the efficacy of interventions, but how they are planned, conducted, and reported raises important concerns.^{5,6} Empirical evidence shows that using inadequate methods to generate randomisation sequences, not concealing allocation, lack of blinding, and excluding patients from analyses can bias findings.⁸⁻¹² Yet half of RCTs fail to take adequate steps to reduce such bias.³⁻¹³

Poor reporting of methods is another common problem.^{3,5} Although it does not necessarily reflect poor methods,^{14,15} poor reporting prevents readers from adequately assessing whether the methods are reliable and whether the results and conclusions of RCTs can be trusted. It also limits reproducibility.¹⁶ RCTs that use inadequate methods or that are poorly reported might not contribute to the evidence base, which is a waste of resources that affects not only RCTs but also the systematic reviews that include them.

Many methodological studies have evaluated the quality of reporting and risk of bias in RCTs.¹⁷⁻³⁰ But most have focused on small numbers of trials or specific diseases, journals, or time periods. These studies used various criteria for their assessments, which were frequently not defined.¹⁷ We therefore lack knowledge of the global magnitude of poor reporting and inadequate methods in RCTs and how they have changed over time. A comprehensive picture of the quality of research would help us better understand the disparities between journals and to propose practical ways of improvement.

Cochrane reviews are uniquely placed to evaluate the quality of research. They synthesise findings from RCTs and are used to aid decision making and develop practice guidelines.³¹⁻³² The risk of bias of each included study is systematically assessed in duplicate by trained reviewers who reach consensus. Reviewers use the Cochrane risk of bias tool, which assesses several methodological items that are considered crucial to evaluating the validity of an RCT.³³⁻³⁴

We examined how poor reporting and inadequate methods have changed over the past three decades, both in and between journals, using the data included in Cochrane reviews.

Methods

We used the proportion of trials considered by the review authors to be at high or unclear risk of bias as surrogate measures of inadequate methods and poor reporting, respectively. We focused on several key items of the Cochrane risk of bias tool: sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessors, and incomplete outcome data (table 1) because they are associated with intervention effect estimates in meta-epidemiological studies.³³⁻³⁶

Data sources

We obtained data from all Cochrane reviews published between March 2011 and September 2014. We chose March 2011 because it corresponded to the most recent update of the Cochrane risk of bias tool.³⁴ An XML file was provided for each review consisting of all data entered by the review authors in Review Manager, the software used for preparing and maintaining Cochrane reviews.³⁷ Each file contained information for all studies included in the review, including methods summary, references, and consensus on risk of bias assessment between the reviewers.³⁴

We downloaded the Web of Science and PubMed databases for the list of indexed journals and the *Journal Citation Reports* database for the list of journals with their medical categories and 2014 impact factors.

Constitution of a unique database

We combined all individual XML files into a single database using R v3.2.2 with the XML package (<https://www.R-project.org/>). We first standardised the wording about risk of bias items because it varied across reviews. Two of us (AD, LT) manually classified them using a standardised set of risk of bias items, and all disagreements were resolved by discussion.

Selection of eligible Cochrane reviews

If the Cochrane review had been updated within our search limits, we considered the most recent version. We excluded reviews that had been withdrawn and “empty” reviews (that is, those not including any studies). We focused on reviews of RCTs only and excluded those of observational or

non-randomised studies. To identify observational or non-randomised studies, two of us (AD, LT) made a list of keywords that could correspond to observational studies—for example, “observational,” “cohort,” and “case-control”—that were automatically searched in the free text description of the methods summary. We also identified risk of bias items that could correspond to observational studies; for example, “potential confounding factors taken into account.” We excluded reviews reporting these keywords or items.

Then, we identified reviews that included an assessment of the risk of bias for the following key items: sequence generation, allocation concealment, blinding (whatever the type of blinding item), and incomplete outcome data. Some reviews assessed blinding overall (according to a previous version of the risk of bias tool), with no distinction between blinding of participants and personnel and blinding of outcome assessors. So we considered reviews eligible if they reported at least one item concerning blinding—blinding overall, blinding of participants and personnel, or blinding of outcome assessors.

Selection of eligible RCTs and identification of corresponding primary reference

We excluded RCTs if the results were not reported in at least one journal article published after 1985. We excluded RCT results published before 1986 (corresponding to the 10th centile of trial year of publication) to focus on contemporary trials. We first extracted the references for all RCTs included in the eligible Cochrane reviews, including the reference type (journal article, book section, or conference proceedings). Then we used a matching algorithm³⁸ and manual validation to identify duplicate references, and we excluded trials that were included in more than one review. For RCTs that referenced more than one journal article, we selected the primary reference reported by the review authors. For the few cases that had several primary references, we manually identified the reference corresponding to reporting of the main results. We excluded RCTs for which no primary reference was reported. We also excluded RCTs reported in abstract format only (reported as such in the characteristics of included studies). We extracted the year of publication and journal names of the primary reference for all selected RCTs.

Matching journal names with Web of Science, PubMed, and Journal Citation Reports

We used the Web of Science and PubMed databases to standardise journal names and abbreviations. One of us (AD) manually reviewed journal names that could not be matched to verify whether they corresponded to existing journals that were not indexed. This enabled us to identify variations in journal names (for example, *Critical Care*, *Critical Care London*, and *Critical Care London England*), which were corrected according to Web of Science or PubMed. We excluded RCTs with journal names corresponding to non-existing journals. Finally, for each journal we extracted the 2014 journal impact factor, medical category, average impact factor centile across medical categories, country, and language from *Journal Citation Reports*. For journals not included in *Journal Citation Reports*, we manually evaluated the main medical category and the language.

Extraction of risk of bias assessment

For each RCT we extracted the Cochrane risk of bias judgments, which were “low,” “high,” or “unclear” for each of the five key items (sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessors, and incomplete outcome data). When a Cochrane review

assigned more than one judgment to the same item (for example, blinding might have been assessed for each outcome) we considered the risk of bias judgment corresponding to subjective outcomes, because objective outcomes should be low risk. For trials that assessed blinding overall, we considered the risk of bias to be the same for both blinding of participants and personnel and blinding of outcome assessors. For trials that were included in more than one Cochrane review, we extracted the risk of bias assessment for the most recent.

Summary of data available for each included RCT

Publication characteristics—year of publication, journal name, medical category, language, and whether the journal was indexed in Web of Science, PubMed, and *Journal Citation Reports*. For RCTs published in a journal indexed in *Journal Citation Reports*, we also had the journal 2014 impact factor.

Risk of bias assessment—final judgment, corresponding to the consensus of two trained reviewers, for: sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessors, and incomplete outcome data.

Assessment of poor reporting and inadequate methods

The Cochrane Handbook says that risk of bias should be considered unclear if there is insufficient information to permit judgment of low or high risk.^{33 34} So an item classified as unclear can be considered poorly reported. A judgment of high risk is given when inadequate methods or conduct could result in a bias of sufficient magnitude to have a notable effect on the results or conclusions of the trial.³⁴ We used the proportion of trials considered by the review authors to be at unclear or high risk of bias as surrogates for poor reporting and inadequate methods, respectively.

Analysis

The analysis was descriptive. We first assessed the overall proportion of trials at unclear and high risk of bias. Then we examined how poor reporting and inadequate methods changed over time. Because inadequate methods can only be assessed when reporting is adequate, we performed two analyses for the evolution of inadequate methods: one based on all trials and one based on only trials that were not at unclear risk.

We predefined four subgroup analyses according to: impact factor (≥ 10 , 5–10, < 5 , no IF); centile of impact factor within the medical categories (≥ 90 th centile, 70–90th, < 70 th percentile, no IF); medical subject category according to *Journal Citation Reports* (“medicine, general and internal” versus other); and each of the 10 journals with the most RCTs. One reviewer suggested an additional subgroup analysis on language (English only versus other).

Patient involvement

Patients were not involved in any aspect of the study design, conduct, or in the development of the research question or outcome measures. There are no plans to disseminate the results of the research to study participants or the relevant patient community. This study is research on existing published research and therefore there was no active patient recruitment for data collection.

Results

Selection and general characteristics

The selection process is reported in web appendix 1. We included data from 2001 systematic reviews, including 20 920 unique articles (median year of publication 2003, interquartile range 1997–2008) published in 3136 journals. 19 551 (93.4%) articles were published in 2390 journals indexed in Web of Science or PubMed, and 17 944 (85.8%) were published in 1706 journals indexed in *Journal Citation Reports*. The median impact factor for the journals was 3.4 (interquartile range 2.0–5.5) (table 2). Most RCTs were published in journals with impact factors < 5 or between 5 and 10 (12 496 (59.7%) and 3134 (15.0%), respectively). Of the journals indexed in *Journal Citation Reports*, 165 (9.7%) were non-English and included 584 RCTs, corresponding to 3.2% of RCTs published in journals with impact factors. 2976 RCTs were published in 1430 journals without impact factors (14.2%).

Characteristics of journals without impact factors

Among the 1430 journals without impact factors, 743 (52.0%) were non-English and included 1511 RCTs (50.8% of RCTs published in journals without impact factors). The most common medical categories were integrative and complementary medicine and medicine, general and internal with 521 RCTs for each. We identified 12 journals (including 29 RCTs) that did not have an impact factor in 2014 but had one in 2015 and 102 journals (including 287 RCTs) that had had an impact factor before 2014.

Overall assessment of poor reporting and inadequate methods

The proportion of trials at unclear risk of bias was high for sequence generation and allocation concealment (48.7% and 57.5%, respectively) and lower for blinding of participants and personnel and incomplete outcome data (30.6% and 24.7%, respectively) (fig 1). The proportion of trials at high risk of bias was 4.0% and 7.2% of all trials for sequence generation and allocation concealment, respectively, but was 33.1% and 22.6% for blinding of participants and personnel and blinding of outcome assessors, respectively, and 17.1% for incomplete outcome data (fig 1).

For all five risk of bias items, we found a lower proportion of trials at unclear or high risk of bias in journals with high impact factors than in those with low or no impact factor (fig 2). For allocation concealment, for example, 38.0% of trials published in journals with impact factors ≥ 10 were at unclear risk of bias, compared with 73.4% of those in journals with no impact factor. We found the same trend when grouping the journals by centiles of impact factor—45.9% for trials published in journals with impact factors above the 90th centile v 73.4% in those with no impact factor. The proportion of trials at high or unclear risk of bias was lower for journals in the “medicine, general and internal” category than for those in other categories and those not indexed (fig 2). We describe risk of bias for the 10 main medical categories in web appendix 2.

A lower proportion of trials were at unclear or high risk of bias in English journals than in non-English journals (web appendix 3). For allocation concealment, 55.7% and 6.6% of RCTs were at unclear and high risk of bias in English journals, compared with 73.7% and 12.5%, respectively, in non-English journals.

Evolution of poor reporting and inadequate methods over time

Evolution of poor reporting

The proportion of trials at unclear risk of bias decreased over time, especially for sequence generation, which fell from 69.1% in 1986-1990 to 31.2% in 2011-14, and for allocation concealment, which fell from 70.1% to 44.6% in the same period (fig 3).

The fall in unclear risk of bias over time for sequence generation and allocation concealment was consistent across all types of journals but seemed more marked for journals with higher impact factors (fig 4) and for general journals compared with specialist journals (web appendix 4).

The evolution of poor reporting for the 10 journals with the most RCTs is shown in web appendix 5. All journals showed an improvement over time but with differences between journals.

Evolution of inadequate methods

Evolution of inadequate methods over time is shown in fig 5. When considering all trials, including those considered at unclear risk of bias, the change seems minimal. For sequence generation, the proportion of trials at high risk of bias dropped from 4.6% in 1986-1990 to 3.2% in 2011-14; for allocation concealment, the proportion fell from 9.8% to 6.4% over the same period. The decrease was greater after excluding trials at unclear risk of bias, from 14.8% to 4.6% and from 32.7% to 11.6%, respectively. We found a slight decrease for blinding of outcome assessors and incomplete outcome data, from 24.0% to 20.3% and from 19.8% to 14.5%, respectively, when considering all trials and from 42.0% to 30.6% and from 28.3% to 18.7% after excluding trials at unclear risk of bias. By contrast, the proportion of trials at high risk of bias for blinding of participants and personnel slightly increased from 31.0% to 36.1% for all trials and from 47.3% to 49.3% after excluding trials at unclear risk of bias. We found no clear difference in evolution over time by journal impact factor (web appendix 6).

Discussion

We extensively mapped the research included in Cochrane reviews and found a fall in poor reporting over time, especially for sequence generation and allocation concealment. The proportion of trials with inadequate methods has also decreased slightly for these items. But we found important differences between journals based on their impact factor and between general and specialist journals. Our results raise concerns about trials published in journals without impact factors, in light of their high number (2976 trials; 14.2% of our sample) and the prevalence of poor reporting and inadequate methods.

Strengths and weaknesses

We built a comprehensive database of primary research (RCTs) by compiling a large amount of data routinely collected for Cochrane systematic reviews. These data are of good quality,³⁹ are standardised in part, and are available in an electronic format. Using these data, we identified more than 20 000 RCTs and the corresponding risk of bias assessments, which were collected in duplicate by trained Cochrane reviewers. Although RCTs included in Cochrane reviews do not represent all RCTs, they cover a large and important body of evidence.^{31 32} We think that research on such a large group of trials would not have been possible without Cochrane reviews.

Our study has several limitations. We relied on Cochrane reviewers' assessments of risk of bias. Although reviewers should be trained in use of the risk of bias tool, variability might exist. For RCTs included in more than one review, we relied on the risk of bias assessment in the most recent. We compared this assessment with previous reviews for 1065 RCTs that shared the same primary reference and found agreement in 83% (n=881) and 75% (n=802) of RCTs for sequence generation and allocation concealment, respectively. Most disagreements concerned the distinction between low and unclear risk of bias. Some Cochrane reviewers might contact study authors for clarification and additional information for some methodological elements that were not clearly reported in study reports. This variability seems to be random and does not seem to be correlated with the time of conduct of the review.

Reviewers might have excluded some trials because of inadequate methods, leading us to underestimate the number of studies with poor reporting and inadequate methods in publications. We cannot exclude the possibility that a classification bias might explain the differences by impact factor. Cochrane reviewers might be influenced by the impact factor of the journal when assessing the risk of bias, with attribution of better scores to journals with the highest impact factors. We think that this is more likely to affect the extreme categories (impact factors ≥ 10 and no impact factor), and we observed a clear trend for all categories.

We relied on the 2014 impact factor from *Journal Citation Reports*, which could differ from that in the year the trial was published. Finally, we focused on inadequate methods and poor reporting and did not consider other important sources of waste in RCTs, such as research questions not relevant to patients or their doctors or failure to report trial results.^{3 7}

Comparison with other studies

This study goes beyond previous literature on this topic with assessment of changes over time and comparisons between journals. Our results show the magnitude of poor reporting in RCTs included in Cochrane reviews, with around half of these RCTs considered at unclear risk of bias by the review authors for sequence generation and allocation concealment, which agrees with previous findings.¹³⁻⁴¹ We found an improvement in reporting for these items over time, with the proportion of trials at unclear risk of bias being halved over three decades, which is consistent with the results of a study of variation in risk of bias over time from a sample of 1732 RCTs included in 97 systematic reviews, published in 2012.⁴⁰

We found a lower proportion of RCTs with poor reporting and inadequate methods in journals with higher impact factors. General medical journals seemed to be associated with lower risk of bias and better reporting than specialist journals, but this may be explained by the higher impact factors of general medical journals.

Journal impact factor might be a surrogate for other factors. Previous studies show that trials published in journals with high impact factors are more likely to report methodological safeguards against bias^{42 43} and to adhere to reporting guidelines than those with lower impact factors.^{43 44} Journals with high impact factors might have more technical resources, which can help to ensure adherence to reporting guidelines by checking submission of the checklist or might be used to detect selective reporting of outcomes by checking information from clinical trial registries. Moreover, journals with high impact factors might be more engaged in quality improvement, with methodologists involved in peer review. Finally, the selection

of trials to be published might be more stringent in journals with high impact factors. However, although some studies support a relationship between impact factor and quality,⁴²⁻⁴⁴ others show no such relationship,⁴⁵ and impact factors remain controversial.⁴⁶⁻⁴⁷ Publication in a journal with a high impact factor does not ensure that an RCT is at low risk of bias.⁴²

Trials published in a journal without an impact factor represented 14.2% of all trials included in Cochrane reviews we examined. Poor reporting and inadequate methods were particularly common in these trials, with limited improvement over time. Although impact factors should not be assumed to represent journal quality, journals without an impact factor are likely to be different from those that do have one. In our sample we identified 1430 journals without impact factors, 52.0% of which were non-English journals (compared with 9.7% of journals with impact factors).

Implications

The improvements over time that we observed are encouraging but could be better. Most of the waste related to poor reporting or inadequate methods could be avoided. We previously showed that half of the waste related to using inadequate methods could be limited at the planning stage of the trial with simple and inexpensive methodological adjustments.¹³ Waste related to poor reporting could be completely avoided. Although using reporting guidelines, such as the CONSORT checklist, is associated with more complete reporting,⁴⁸⁻⁴⁹ their implementation varies between journals⁴⁴⁻⁵² with many journals having no policy or mentioning only the existence of the CONSORT statement in the instructions to authors.⁴⁴⁻⁵² We need to promote more active implementation, such as submission of the checklist with the manuscript, as it has been associated with better reporting.⁵¹

Research investigators and other stakeholders must also act responsibly to report clear, transparent, and reliable research findings.⁶ This responsibility can be communicated through education and training starting at university and continuing through residency, fellowship, and the various stages of a career (for example, through professional societies and university departments). Investigators should work with methodologists from the planning stage of their trial to increase the likelihood of adequate study design and quality of reporting.⁵³ If having a methodologist as a co-investigators is not possible, writing aid tools might be useful.⁵⁴

Our data provide an overview of the quality of evidence in Cochrane reviews. Poor reporting and inadequate methods are common in RCTs included in Cochrane reviews and might affect their results and conclusions, as previous meta-epidemiological studies have shown.⁸⁻¹² Our results highlight the importance of assessing risk of bias and of incorporating this assessment in evidence synthesis. Improvements at the trial level are necessary to improve the quality of evidence provided by systematic reviews.

This study is part of a larger project, the next step of which is to talk to journals either alone or perhaps grouped by specialty⁵⁵ about whether this type of audit and feedback is useful to their journal practice. Such data can serve an important monitoring function, examining incremental changes in quality over time. If our process could be automated, this might be a powerful tool for funding agencies and others interested in assessing the value of their research investments.¹⁻⁵⁶ Cochrane is uniquely placed to observe the quality of research, as its data are routinely collected with excellent quality assurance.

Conclusion

This extensive mapping of trials shows a decrease in waste related to poor reporting and inadequate methods over time, but with important differences between risk of bias items and between journals. Our approach, based on the use of data already collected by Cochrane for its reviews, could be a first step in the development of a live observatory to monitor the quality of research over time.

We thank David Tovey, editor in chief of the Cochrane Library, for sharing data from Cochrane reviews; Javier Mayoral Campos, system administrator; the Cochrane Central Executive for preparing files; and all Cochrane reviewers who collected data. We also thank Carolina Riveros, from Inserm U1153, for help with data extraction and Elise Diard, from Inserm U1153, for help with figures.

Contributors: PR generated the idea; AD, LT, and PR designed the study; AD, LT, and IA selected and collected the data; IA and EP managed the data and conducted the statistical analysis; AD, LT, DM, KD, IB, DGA, PR interpreted the data; AD and PR wrote the manuscript; and LT, IA, DM, KD, IB, and DGA critically reviewed the manuscript. AD is the guarantor. She had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Funding: The researchers did not receive external funding. This work was internally funded by Cochrane France. The funding source had no role in the design, conduct, writing, or submission of this article.

Competing interests: DM is a member of the Cochrane Library Oversight Committee and a member of the Cochrane Bias Methods Group. He received funding from the Cochrane Collaboration for an unrelated project. IB and DGA are also members of the Cochrane Bias Methods Group. DGA is supported by Cancer Research UK (C5529). KD is the director of the US Cochrane Center. PR is the director of Cochrane France. The other authors declare no competing interests.

Ethical approval: Not applicable. This is a research on research study.

Data sharing: Data and analysis code available on request from the authors.

Transparency declaration: The guarantor (AD) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

- 1 Al-Shahi Salman R, Beller E, Kagan J, et al. Increasing value and reducing waste in biomedical research regulation and management. *Lancet* 2014;357:176-85. doi:10.1016/S0140-6736(13)62297-7 pmid:24411646.
- 2 Chalmers I, Bracken MB, Djulibegovic B, et al. How to increase value and reduce waste when research priorities are set. *Lancet* 2014;357:156-65. doi:10.1016/S0140-6736(13)62229-1 pmid:24411644.
- 3 Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009;357:86-9. doi:10.1016/S0140-6736(09)60329-9 pmid:19525005.
- 4 Chan AW, Song F, Vickers A, et al. Increasing value and reducing waste: addressing inaccessible research. *Lancet* 2014;357:257-66. doi:10.1016/S0140-6736(13)62296-5 pmid:24411650.
- 5 Glasziou P, Altman DG, Bossuyt P, et al. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 2014;357:267-76. doi:10.1016/S0140-6736(13)62228-X pmid:24411647.
- 6 Ioannidis JP, Greenland S, Hlatky MA, et al. Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014;357:166-75. doi:10.1016/S0140-6736(13)62227-8 pmid:24411645.
- 7 Macleod MR, Michie S, Roberts I, et al. Biomedical research: increasing value, reducing waste. *Lancet* 2014;357:101-4. doi:10.1016/S0140-6736(13)62329-6 pmid:24411643.
- 8 Moher D, Pham B, Jones A, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998;357:609-13. doi:10.1016/S0140-6736(98)01085-X pmid:9746022.
- 9 Nüesch E, Trelle S, Reichenbach S, et al. The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study. *BMJ* 2009;357:b3244. doi:10.1136/bmj.b3244 pmid:19736281.
- 10 Pildal J, Hróbjartsson A, Jørgensen KJ, Hilden J, Altman DG, Gotzsche PC. Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. *Int J Epidemiol* 2007;357:847-57. doi:10.1093/ije/dym087 pmid:17517809.
- 11 Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;357:408-12. doi:10.1001/jama.1995.03520290060030 pmid:7823387.

What is already known on this subject

Poor reporting and inadequate methods are common in randomised controlled trials (RCTs)

Many methodological studies have evaluated the quality of reporting and risk of bias in RCTs, but they are limited in terms of number of trials evaluated, most focusing on specific diseases, journals, or time periods.

What this study adds

We took advantage of the amount and quality of data included in Cochrane reviews to map the evolution of poor reporting and inadequate methods in and between journals

From nearly 21 000 RCTs published in 3136 journals over three decades, our results show a decrease over time of poor reporting and inadequate methods especially for sequence generation and allocation concealment

We found a lower proportion of RCTs with poor reporting and inadequate methods in journals with higher impact factors. By contrast, our results raise concerns about journals without impact factors because of the prevalence of poor reporting and inadequate methods

The next step would be to provide feedback to journals and to evaluate whether this type of audit has an impact by using Cochrane data as a live observatory to monitor changes over time

- 12 Tierney JF, Stewart LA. Investigating patient exclusion bias in meta-analysis. *Int J Epidemiol* 2005;35:79-87. doi:10.1093/ije/dyh300 pmid:15561753.
- 13 Yordanov Y, Dechartres A, Porcher R, Boutron I, Altman DG, Ravaud P. Avoidable waste of research related to inadequate methods in clinical trials. *BMJ* 2015;357:h809. doi:10.1136/bmj.h809 pmid:25804210.
- 14 Mhaskar R, Djulbegovic B, Magazín A, Soares HP, Kumar A. Published methodological quality of randomized controlled trials does not reflect the actual quality assessed in protocols. *J Clin Epidemiol* 2012;357:602-9. doi:10.1016/j.jclinepi.2011.10.016 pmid:22424985.
- 15 Soares HP, Daniels S, Kumar A, et al. Radiation Therapy Oncology Group. Bad reporting does not mean bad methods for randomised trials: observational study of randomised controlled trials performed by the Radiation Therapy Oncology Group. *BMJ* 2004;357:22-4. doi:10.1136/bmj.328.7430.22 pmid:14703540.
- 16 Altman DG. Making research articles fit for purpose: structured reporting of key methods and findings. *Trials* 2015;357:53. doi:10.1186/s13063-015-0575-7 pmid:25888056.
- 17 Dechartres A, Charles P, Hopewell S, Ravaud P, Altman DG. Reviews assessing the quality or the reporting of randomized controlled trials are increasing over time but raised questions about how quality is assessed. *J Clin Epidemiol* 2011;357:136-44. doi:10.1016/j.jclinepi.2010.04.015 pmid:20705426.
- 18 Agha RA, Camm CF, Doganay E, Edison E, Siddiqui MR, Orgill DP. Randomised controlled trials in plastic surgery: a systematic review of reporting quality. *Eur J Plast Surg* 2014;357:55-62. doi:10.1007/s00238-013-0893-5 pmid:24707112.
- 19 Bryant J, Passey ME, Hall AE, Sanson-Fisher RW. A systematic review of the quality of reporting in published smoking cessation trials for pregnant women: an explanation for the evidence-practice gap? *Implement Sci* 2014;357:94. doi:10.1186/s13012-014-0094-z pmid:25138616.
- 20 Chen B, Liu J, Zhang C, Li M. A retrospective survey of quality of reporting on randomized controlled trials of metformin for polycystic ovary syndrome. *Trials* 2014;357:128. doi:10.1186/1745-6215-15-128 pmid:24746168.
- 21 Chen X, Zhai X, Wang X, Su J, Li M. Methodological reporting quality of randomized controlled trials in three spine journals from 2010 to 2012. *Eur Spine J* 2014;357:1606-11. doi:10.1007/s00586-014-3283-1 pmid:24748442.
- 22 Chen Z, Chen Y, Zeng J, et al. Quality of randomized controlled trials reporting in the treatment of melasma conducted in China. *Trials* 2015;357:156. doi:10.1186/s13063-015-0677-2 pmid:25872530.
- 23 Glujovsky D, Boggino C, Riestra B, Coscia A, Sueldo CE, Ciapponi A. Quality of reporting in infertility journals. *Fertil Steril* 2015;357:236-41. doi:10.1016/j.fertnstert.2014.10.024 pmid:25455871.
- 24 Kim KH, Kang JW, Lee MS, Lee JD. Assessment of the quality of reporting in randomised controlled trials of acupuncture in the Korean literature using the CONSORT statement and STRICTA guidelines. *BMJ Open* 2014;357:e005068. doi:10.1136/bmjopen-2014-005068 pmid:25079926.
- 25 Kloukos D, Papageorgiou SN, Doulis I, Petridis H, Pandis N. Reporting quality of randomised controlled trials published in prosthodontic and implantology journals. *J Oral Rehabil* 2015;357:914-25. doi:10.1111/joor.12325 pmid:26132229.
- 26 Lee SY, Teoh PJ, Camm CF, Agha RA. Compliance of randomized controlled trials in trauma surgery with the CONSORT statement. *J Trauma Acute Care Surg* 2013;357:562-72. doi:10.1097/TA.0b013e3182a5399e pmid:24064867.
- 27 Lempesi E, Koletsis D, Fleming PS, Pandis N. The reporting quality of randomized controlled trials in orthodontics. *J Evid Based Dent Pract* 2014;357:46-52. doi:10.1016/j.jebdp.2013.12.001 pmid:24913524.
- 28 Yao AC, Khajuria A, Camm CF, Edison E, Agha R. The reporting quality of parallel randomised controlled trials in ophthalmic surgery in 2011: a systematic review. *Eye (Lond)* 2014;357:1341-9. doi:10.1038/eye.2014.206 pmid:25214001.
- 29 Zhai X, Wang Y, Mu Q, et al. Methodological Reporting Quality of Randomized Controlled Trials in 3 Leading Diabetes Journals From 2011 to 2013 Following CONSORT Statement: A System Review. *Medicine (Baltimore)* 2015;357:e1083. doi:10.1097/MD.0000000000001083 pmid:26166088.
- 30 Zhuang L, He J, Zhuang X, Lu L. Quality of reporting on randomized controlled trials of acupuncture for stroke rehabilitation. *BMC Complement Altern Med* 2014;357:151. doi:10.1186/1472-6882-14-151 pmid:24885561.
- 31 Bunn F, Trivedi D, Alderson P, et al. The impact of Cochrane Reviews: a mixed-methods evaluation of outputs from Cochrane Review Groups supported by the National Institute for Health Research [v-vi]. *Health Technol Assess* 2015;357:1-99, v-vi. doi:10.3310/hta19280 pmid:25875129.
- 32 Davoli M, Amato L, Clark N, et al. The role of Cochrane reviews in informing international guidelines: a case study of using the Grading of Recommendations, Assessment, Development and Evaluation system to develop World Health Organization guidelines for the psychosocially assisted pharmacological treatment of opioid dependence. *Addiction* 2015;357:891-8. doi:10.1111/add.12788 pmid:25490943.
- 33 Higgins JP, Altman DG, Gøtzsche PC, et al. Cochrane Bias Methods Group Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;357:d5928. doi:10.1136/bmj.d5928 pmid:22008217.
- 34 Higgins JPT, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration 2011; Available from <http://handbook.cochrane.org/>.
- 35 Dechartres A, Trinquart L, Faber T, Ravaud P. Empirical evaluation of which trial characteristics are associated with treatment effect estimates. *J Clin Epidemiol* 2016;357:24-37. doi:10.1016/j.jclinepi.2016.04.005 pmid:27140444.
- 36 Page MJ, Higgins JP, Clayton G, Sterne JA, Hróbjartsson A, Savović J. Empirical Evidence of Study Design Biases in Randomized Trials: Systematic Review of Meta-Epidemiological Studies. *PLoS One* 2016;357:e0159267. doi:10.1371/journal.pone.0159267 pmid:27398997.
- 37 Manager R. (RevMan) [Computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.
- 38 deMelo VV. Conference: Advances in Logic Based Intelligent Systems. 2005.
- 39 Page MJ, Shamseer L, Altman DG, et al. Epidemiology and Reporting Characteristics of Systematic Reviews of Biomedical Research: A Cross-Sectional Study. *PLoS Med* 2016;357:e1002028. doi:10.1371/journal.pmed.1002028 pmid:27218655.
- 40 Reveliz L, Chapman E, Asial S, Munoz S, Bonfill X, Alonso-Coello P. Risk of bias of randomized trials over time. *J Clin Epidemiol* 2015;357:1036-45. doi:10.1016/j.jclinepi.2014.06.001 pmid:26227423.
- 41 Savović J, Jones H, Altman D, et al. Influence of reported study design characteristics on intervention effect estimates from randomised controlled trials: combined analysis of meta-epidemiological studies. *Health Technol Assess* 2012;357:1-82. doi:10.3310/hta16350 pmid:22989478.
- 42 Bala MM, Akl EA, Sun X, et al. Randomized trials published in higher vs. lower impact journals differ in design, conduct, and analysis. *J Clin Epidemiol* 2013;357:286-95. doi:10.1016/j.jclinepi.2012.10.005 pmid:23347852.
- 43 Lee KP, Schotland M, Bacchetti P, Bero LA. Association of journal quality indicators with methodological quality of clinical research articles. *JAMA* 2002;357:2805-8. doi:10.1001/jama.287.21.2805 pmid:12038918.
- 44 Samaan Z, Mbuagbaw L, Kosa D, et al. A systematic scoping review of adherence to reporting guidelines in health care literature. *J Multidiscip Healthc* 2013;357:169-88. pmid:23671390.
- 45 Macleod MR, Lawson McLean A, Kyriakopoulou A, et al. Risk of Bias in Reports of In Vivo Research: A Focus for Improvement. *PLoS Biol* 2015;357:e1002273. doi:10.1371/journal.pbio.1002273 pmid:26460723.
- 46 Callaway E. Beat it, impact factor! Publishing elite turns against controversial metric. *Nature* 2016;357:210-1. doi:10.1038/nature.2016.20224 pmid:27411614.
- 47 Van Noorden R. Controversial impact factor gets a heavyweight rival. *Nature* 2016;357:325-6. doi:10.1038/nature.2016.21131 pmid:27974784.
- 48 Cobo E, Cortés J, Ribera JM, et al. Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: masked randomised trial. *BMJ* 2011;357:d6783. doi:10.1136/bmj.d6783 pmid:22108262.
- 49 Turner L, Shamseer L, Altman DG, et al. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. *Cochrane Database Syst Rev* 2012;357:MR000030. pmid:23152285.
- 50 Hopewell S, Altman DG, Moher D, Schulz KF. Endorsement of the CONSORT Statement by high impact factor medical journals: a survey of journal editors and journal 'Instructions to Authors'. *Trials* 2008;357:20. doi:10.1186/1745-6215-9-20 pmid:18423021.
- 51 Hopewell S, Ravaud P, Baron G, Boutron I. Effect of editors' implementation of CONSORT guidelines on the reporting of abstracts in high impact medical journals: interrupted time series analysis. *BMJ* 2012;357:e4178. doi:10.1136/bmj.e4178 pmid:22730543.
- 52 Shamseer L, Hopewell S, Altman DG, Moher D, Schulz KF. Update on the endorsement of CONSORT by high impact factor journals: a survey of journal "Instructions to Authors" in 2014. *Trials* 2016;357:301. doi:10.1186/s13063-016-1408-z pmid:27343072.
- 53 Delgado-Rodríguez M, Ruiz-Canela M, De Irala-Estevez J, Llorca J, Martínez-González A. Participation of epidemiologists and/or biostatisticians and methodological quality of published controlled clinical trials. *J Epidemiol Community Health* 2001;357:569-72. doi:10.1136/jech.55.8.569 pmid:11449014.
- 54 Barnes C, Boutron I, Giraudeau B, Porcher R, Altman DG, Ravaud P. Impact of an online writing aid tool for writing a randomized trial report: the COBWEB (Consort-based WEB tool) randomized controlled trial. *BMC Med* 2015;357:221. doi:10.1186/s12916-015-0460-y pmid:26370288.
- 55 Chan L, Heinemann AW, Roberts J. Elevating the quality of disability and rehabilitation research: mandatory use of the reporting guidelines. *Arch Phys Med Rehabil* 2014;357:415-7. doi:10.1016/j.apmr.2013.12.010 pmid:24559651.
- 56 Moher D, Glasziou P, Chalmers I, et al. Increasing value and reducing waste in biomedical research: who's listening? *Lancet* 2016;357:1573-86. doi:10.1016/S0140-6736(15)00307-4 pmid:26423180.

Accepted: 09 05 2017

Published by the BMJ Publishing Group Limited. For permission to use (where not already granted under a licence) please go to <http://group.bmj.com/group/rights-licensing/>

permissions

Tables

Table 1| Items in the Cochrane risk of bias tool

Item	Definition	Types of bias
Sequence generation	Method used to generate the allocation sequence reported in sufficient detail to enable assessment of whether it should produce comparable groups	Selection bias (biased allocation to interventions) due to inadequate generation of a randomised sequence
Allocation concealment	Method used to conceal the allocation sequence reported in sufficient detail to determine whether intervention allocations could have been foreseen before, or during, enrolment	Selection bias (biased allocation to interventions) due to inadequate concealment of allocations before assignment
Blinding of participants and personnel	Measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received	Performance bias due to knowledge of the allocated interventions by participants and personnel during the study
Blinding of outcome assessors	Measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received	Detection bias due to knowledge of the allocated interventions by outcome assessment
Incomplete outcome data	Evaluation of the completeness of outcome data for each main outcome, including attrition and number of exclusions from the analysis	Attrition bias due to amount, nature, or handling of incomplete outcome data

Sources: Cochrane Handbook³⁴ and Higgins JP et al.³³

Table 2| General characteristics of included RCTs (n=20 920)

Characteristic	n (%) unless otherwise specified
Publication year*	2003 (1997-2008), 1986-2014
Journals indexed in PubMed or Web of Science	19 551 (93.4)
Journals indexed in JCR in the following categories:†	17 944 (85.8)
Medicine: general and internal	2204 (12.3)
Obstetrics and gynaecology	2000 (11.1)
Paediatrics	1347 (7.5)
Surgery	1337 (7.4)
Clinical neurology	1234 (6.9)
Psychiatry	1161 (6.5)
Pharmacology and pharmacy	897 (5.0)
Anaesthesiology	796 (4.4)
Oncology	775 (4.3)
Respiratory system	711 (4.0)
Other	6732 (37.5)
Published in journal with impact factor	17 944 (85.8)
Journal impact factor:*	3.4 (2.0-5.5), 0.05-55.9
≥10	2314 (11.1)
5-10	3134 (15.0)
<5	12 496 (59.7)
Published in journal without an impact factor	2976 (14.2)
10 highest represented journals:‡	2386 (11.4)
<i>New England Journal of Medicine</i>	384 (1.8)
<i>The Lancet</i>	367 (1.7)
<i>American Journal of Obstetrics and Gynecology</i>	257 (1.2)
<i>Obstetrics and Gynaecology</i>	225 (1.1)
<i>The BMJ</i>	224 (1.1)
<i>BJOG: An International Journal of Obstetrics and Gynaecology</i>	219 (1.0)
<i>Pediatrics</i>	207 (1.0)
<i>Journal of the American Medical Association</i>	194 (0.9)
<i>Journal of Pediatrics</i>	158 (0.8)
<i>Journal of Clinical Oncology</i>	151 (0.7)

*median (interquartile range), min-max

†A journal can have several categories according to the Journal Citation Reports

‡80% of RCTs were published in 24% of the journals. The 10 journals with the most RCTs included 2386 trials, representing 11% of all RCTs, and comprised four general medical journals (*New England Journal of Medicine*, *The Lancet*, *The BMJ*, and *Journal of the American Medical Association*) and six specialist journals (*American Journal of Obstetrics and Gynaecology*, *Obstetrics and Gynaecology*, *BJOG: An International Journal of Obstetrics and Gynaecology*, *Paediatrics*, *Journal of Pediatrics*, and *Journal of Clinical Oncology*).

Figures

Fig 1 Risk of bias for each key methodological item in 20 920 trial articles

Fig 2 Risk of bias for each key methodological item in 20 920 trial articles by journal impact factor and medical category. IF=impact factor.

Fig 3 Evolution of poor reporting over time in 20 920 trial articles. The proportion of trials at unclear risk of bias is represented as a surrogate for poor reporting. Data are proportions and 95% confidence intervals

Fig 4 Evolution of poor reporting over time in 20 920 trial articles by journal impact factor. The proportion of trials at unclear risk of bias is represented as a surrogate for poor reporting. IF=impact factor.

Fig 5 Evolution of inadequate methods over time for a) all trials and b) trials not at unclear risk of bias for the item considered. The proportion of trials at high risk of bias is represented as a surrogate for inadequate methods. Data are proportions and 95% confidence intervals