

RESEARCH

Open Access



Opportunities for selective reporting of harms in randomized clinical trials: Selection criteria for non-systematic adverse events

Evan Mayo-Wilson^{1*†} , Nicole Fusco^{1†}, Hwanhee Hong², Tianjing Li¹, Joseph K. Canner³ and Kay Dickersin¹

Abstract

Background: Adverse events (AEs) in clinical trials may be reported in multiple sources. Different methods for reporting adverse events across trials or across sources for a single trial may produce inconsistent information about the adverse events associated with interventions.

Methods: We compared the methods authors use to decide which AEs to include in a particular source (i.e., “selection criteria”), including the number of different types of AEs reported (i.e., rather than the number of events). We compared sources (e.g., journal articles, clinical study reports (CSRs)) of trials for two drug-indications—gabapentin for neuropathic pain and quetiapine for bipolar depression. Electronic searches were completed in 2015. We identified selection criteria and assessed how criteria affected AE reporting.

Results: We identified 21 gabapentin and 7 quetiapine trials. We found 6 gabapentin CSRs and 2 quetiapine CSRs, all written by drug manufacturers. All CSRs reported all AEs without applying selection criteria; by comparison, no other source reported all AEs, and 15/68 (22%) gabapentin sources and 19/48 (40%) quetiapine sources reported using selection criteria. Selection criteria greatly affected the number of AEs reported. For example, 67/316 (21%) AEs in one quetiapine trial met the criterion “occurring in $\geq 2\%$ of participants in any treatment group,” while only 5/316 (2%) AEs met the criterion “occurring in $\geq 10\%$ of quetiapine-treated patients and twice as frequent in the quetiapine group as the placebo group.”

Conclusions: Selection criteria for reporting AEs vary across trials and across sources for individual trials. If investigators do not pre-specify selection criteria, they might “cherry-pick” AEs based on results. Even if investigators pre-specify selection criteria, selective reporting will produce biased meta-analyses and clinical practice guidelines. Data about all AEs identified in clinical trials should be publicly available; however, sharing data will not solve all the problems identified in this study.

Keywords: Harms, Adverse events, Clinical trials, Reporting bias, Selective outcome reporting, Data sharing, Trial registration

* Correspondence: emayowil@iu.edu

†Joint first authors: Evan Mayo-Wilson and Nicole Fusco contributed equally to this paper.

¹Department of Epidemiology and Biostatistics, Indiana University School of Public Health-Bloomington, 1025 E 7th St, #179D, Bloomington, IN 47405, USA

Full list of author information is available at the end of the article



Introduction

The U.S. Food and Drug Administration (FDA) and other regulators may approve drugs and biologic agents for marketing when the results of randomized clinical trials indicate that their potential benefits outweigh their potential harms, which are often called “adverse events” (AEs). In clinical trials, AEs can be collected non-systematically when participants report them spontaneously to investigators or in response to open-ended questions such as “have you noticed any symptoms since your last examination?” (Table 1) [1–4]. Specific AEs can also be collected systematically by assessing their presence or absence using the same method for all participants in a trial (e.g., clinical examination, questionnaires, and medical instruments [1, 2]). Previous studies have described non-systematic AEs as “passively collected” and systematic AEs as “proactively collected” [1, 2]. AEs are categorized as “serious” when they lead to or prolong hospitalization, cause death, or disrupt normal life functions [4, 5].

Physicians, patients, and policy makers rely on systematic reviews and clinical practice guidelines to make medical decisions. Syntheses of clinical trial findings

should include all available evidence; however, they are often based on AE information reported in public sources, such as journal articles (Table 1) [6]. Many systematic reviews that plan to synthesize AEs ultimately do not address AEs [7]. By contrast, regulators have access to non-public sources of trial information. Although previously non-public sources are becoming available for some trials [8–10], including through data sharing services such as Vivli [11], Yale Open Data Access (YODA) [12, 13], and clinicalstudydatarequest.com, clinical study reports (CSRs) and individual patient datasets (IPD) remain unavailable for many trials. Consequently, the results of many trials are available only in public sources [14, 15], which are often incomplete [15–24]. Reporting bias, which is the selective reporting of research results, occurs when reporting is influenced by the nature of the results (e.g., the direction, magnitude, or statistical significance of the results). Reporting bias may lead to overestimating potential benefits of an intervention [15–20] and underestimating potential AEs [22, 25–32].

There is little evidence about the methods trialists use to select AEs for reporting, and there is little

Table 1 Terms related to adverse events (AEs)

Term	Definition
Adverse event (AE)	The International Conference on Harmonisation (ICH) defines an “adverse event” as “any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and which does not necessarily have to have a causal relationship with this treatment” [33]. The U.S. Food and Drug Administration (FDA) and other regulators use this definition [4, 5].
Non-systematic adverse events	According to The Final Rule [1, 2], “non-systematic assessment” relies on the spontaneous reporting of adverse events, such as unprompted self-reporting by participants”. Non-systematic adverse events may be collected by asking questions such as “Have you noticed any symptoms since your last examination?”.
Result	In the Multiple Data Sources (MUDS) study, a “result” is a numerical contrast between a treatment and comparison arm (e.g., relative risk, mean difference).
Serious adverse events	The ICH defines a “serious adverse event” as that which “results in death, is life-threatening, requires inpatient hospitalisation or prolongation of existing hospitalisation, results in persistent or significant disability/incapacity, or is a congenital anomaly/birth defect” [33]. The FDA and other regulators use this definition [4, 5].
Systematic adverse events	According to The Final Rule [1, 2], “systematic assessment” involves the use of a specific method of ascertaining the presence of an adverse event (e.g., the use of checklists, questionnaires, specific laboratory tests at regular intervals). Like a potential benefit of treatment, a systematic AE can be defined using five elements: (1) domain, (2) specific measurement, (3) specific metric, (4) method of aggregation, and (5) time-point [34]. For example, “proportion of participants with 50% change from baseline to 8 weeks on the Young Mania Rating Scale total score.”
Terms related to sources	
Clinical study report (CSR)	A comprehensive document, often created by a pharmaceutical manufacturer for submission to a regulator, detailing the design, methods, analyses, and results of a study. Appendices sometimes contain tables of individual patient data, also called “patient data listings”, and study protocols [35].
Clinical study report synopsis (CSR-synopsis)	A document that summarizes the information contained in a clinical study report. Clinical study report-synopses are much shorter than clinical study reports; the two clinical study report-synopses we examined were each 13 pages in length.
Individual patient data (IPD)	A table or database in which each record contains data for a single participant [35].
Non-public sources	In the MUDS study, non-public sources include individual patient data, clinical study reports, and clinical study report-synopses.
Public sources	In the MUDS study, public sources include journal articles, conference abstracts, commentaries, posters, trial registrations and associated results, and medical reviews and statistical reviews written by the FDA.

evidence about whether those methods contribute to reporting bias. Previous guidance for reporting adverse events has discouraged vague statements about AEs [36, 37], and the Consolidated Standards of Reporting Trials (CONSORT) Extension for Harms discourages authors from reporting AEs above frequency thresholds (e.g., >10%) [38]. Nonetheless, there is little evidence about how different trials and different sources for a single trial report AEs. The objectives of this study were to (1) compare selection criteria for reporting non-systematic AEs (i.e., the methods authors use to decide which AEs to include in a particular source as illustrated in Table 2, (2) examine how different selection criteria could affect AE reporting, and (3) assess how different selection criteria could impact meta-analyses of AEs.

Methods

This analysis is a sub-study of the Multiple Data Sources (MUDS) study. MUDS was a methodologic study, the overall objective of which was to examine multiple data sources about trials and to determine whether the information source could affect the conclusions of systematic reviews and meta-analyses. The published protocol and amendments [24, 39] describe the study methods. Briefly, we searched for public and non-public sources (see definitions in Table 1) and requested additional non-public sources (e.g., CSRs that they had not been disclosed previously) from the companies that manufacture gabapentin and quetiapine as described elsewhere [22, 39, 40]. Electronic searches included the Cochrane Central Register of Controlled Trials (CENTRAL), PubMed, Embase, LILACS, and CINAHL through 2 March 2015 for gabapentin trials and through 26 January 2015 for quetiapine trials. For trial registrations, we searched the International Clinical Trials Registry Platform Search Portal and ClinicalTrials.gov through 10 October 2014.

Table 2 Examples of selection criteria, with each component identified in bold

Two selection criteria (i.e., numerical threshold and participant group):
Adverse events are reported if they occurred in ≥ 5% of participants in any intervention group .
Adverse events are reported if they occurred in ≥ 2% of participants receiving gabapentin .
Three selection criteria:
Adverse events are reported if they occurred in ≥ 2% of participants receiving gabapentin and if they occurred at least twice as frequently in participants receiving gabapentin, compared with participants receiving placebo.

Eligible trials and sources

Eligible studies were parallel randomized clinical trials that compared either gabapentin for neuropathic pain or quetiapine for bipolar depression, with placebo in adults. We excluded open-label and crossover trials. We selected these case studies because we had access to both public and non-public sources for some of the eligible trials.

The MUDS study included 21 gabapentin trials (80 sources, including 6 IPD) and 7 quetiapine trials (52 sources, including 2 IPD). We excluded IPD from this sub-study because IPD did describe methods for analyzing data and because the CSRs included aggregate results that were consistent with the IPD. Thus, we included 68 public sources and 6 non-public sources (all CSRs) for gabapentin trials and we included 46 public sources and 4 non-public sources (2 CSRs, 2 CSR-synopses) for quetiapine trials. All studies were reported in one or more public sources, except one gabapentin trial that was not reported in any public source.

Data extraction

From each source of each eligible trial, two investigators independently extracted data using the open access Systematic Review Data Repository (SRDR; <http://srdr.ahrq.gov/>) and resolved differences by discussion. We extracted results (i.e., the number or proportion of participants who experienced AEs) that were reported for each trial; we did not extract results of analyses that pooled multiple trials.

We classified each AE as “systematic” if its presence or absence was recorded for every participant and assessed using specific measurement tools (e.g., questionnaires, checklists, laboratory tests, or clinical examinations); we classified all other AEs as “non-systematic” (Table 1). For each non-systematic AE, we extracted the name of the event (e.g., “dizziness”, “headache”) and the numerical results closest to 8 and 18 weeks (e.g., proportion of participants in each group who experienced the AE at 6 weeks). We selected 8 weeks (time window 4–13 weeks) a priori as the minimum clinically important time period, and we selected 18 weeks (time window 14–22) because we expected to find some available follow-up data in this time window. We also looked for data at 27 weeks (time window 23–31 weeks) and for longer times, but we did not find trials that reported AEs after the 8 and 18 week time windows.

We extracted the names of non-systematic AEs even when numerical results were not reported; for example, if a source reported that “the most common AEs were dizziness and headache,” we extracted that dizziness and headache were reported and we entered the numerical results as missing data [22].

We extracted the methods from each source that authors said they used to select AEs for inclusion in the

source, which we refer to as “selection criteria” or we recorded that the selection criteria were not reported (Table 2). Although we found evidence that potential benefits and systematic AEs are reported selectively [23], we did not find that systematic AEs were reported based on selection criteria as described in this sub-study. Thus, we describe selection criteria for non-systematic AEs only.

Comparing non-systematic AE selection criteria sources and trials

We compared the selection criteria for non-systematic AEs reported across sources for each trial (see examples in Table 2). We recorded three pre-specified components of non-systematic AE selection criteria:

- Numerical threshold: a cutoff for reporting the number or proportion of participants who experienced a non-systematic AE (e.g., $\geq 5\%$ of participants).
- Participant group: the group(s) in a trial that must have experienced an AE (e.g., participants in the test intervention group, all participants in the trial).
- Difference-between-groups: a cutoff for reporting a difference in the number or proportion of participants who experienced an AE, comparing one participant group with another (e.g., more frequent in the test intervention group compared with the placebo group).

Applying AE selection criteria to data in individual trials

To assess how different selection criteria might affect which and how many non-systematic AEs would be reported, we combined each of the numerical thresholds we observed ($N=5$) with each of the participant-group criteria ($N=3$) and difference-between-groups criteria we observed ($N=3$) to create 45 combined selection criteria.

We observed some selection criteria in both gabapentin and quetiapine trials, and we observed other selection criteria in gabapentin trials only or in quetiapine trials only. Although we analyzed only numerical thresholds, participant groups, and differences between groups that we observed in eligible trials, we never observed some of these combined selection criteria in any trial.

Using all non-systematic AEs reported in each CSR, we then applied each of 45 combined selection criteria to the data found in CSRs for six gabapentin trials and two quetiapine trials. We estimated the number of AEs (i.e., the different types of AEs rather than the number of events) that would have been reported according to each combined criterion.

Statistical methods

We calculated descriptive statistics (i.e., counts) using Stata 14 [41].

Results

AE selection criteria in public and non-public sources

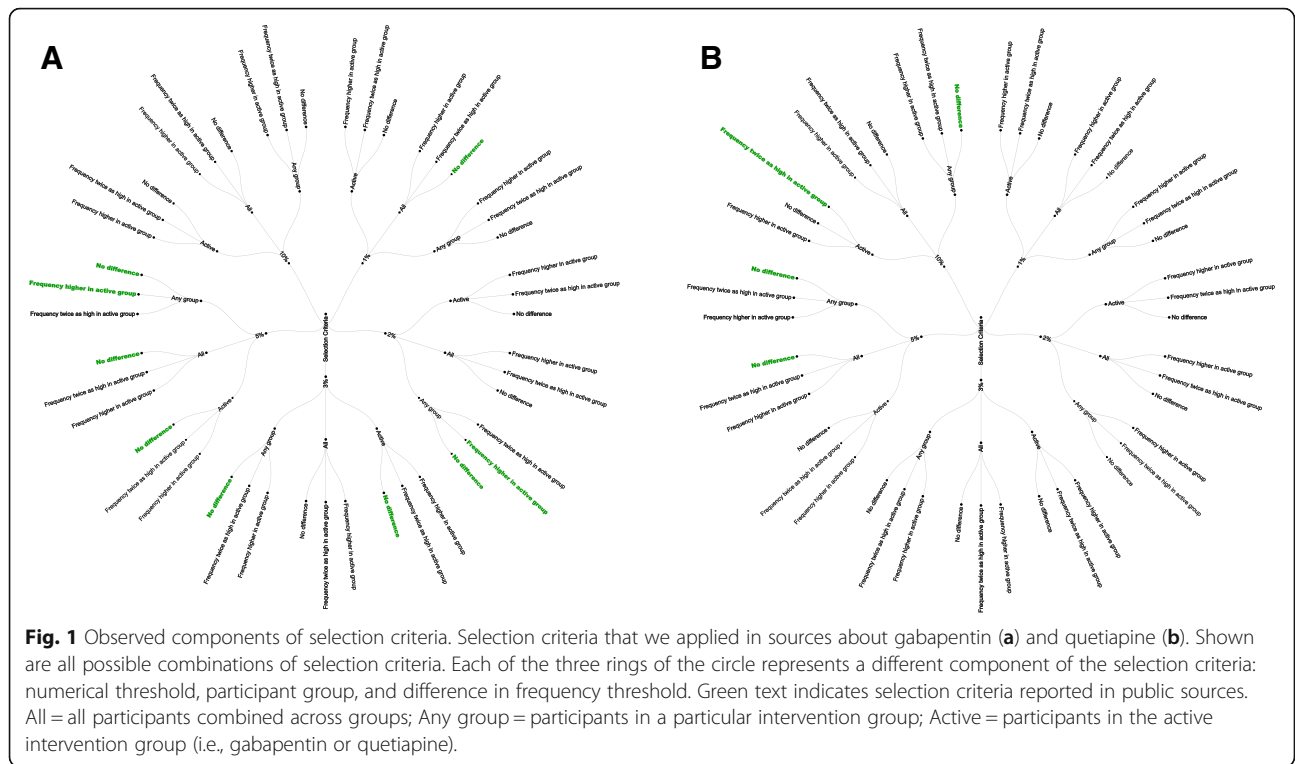
We identified selection criteria only in public sources and in CSR-synopses. CSRs, which included trial protocols, did not report using selection criteria. Based on the information in CSRs, including trial protocols, CSRs appeared to include information about all observed AEs. CSRs did not describe which selection criteria would be used to report AEs in other sources, such as journal articles; thus, we found no scientific rationale or evidence that selection criteria used in other sources were pre-specified [35].

In public sources and CSR-synopses we identified:

- Five different numerical thresholds, including four in gabapentin sources (1%, 2%, 3%, and 5%) and two in quetiapine sources (5% and 10%);
- Three different participant groups; all three participant groups were included in both gabapentin and quetiapine sources. These were (1) participants in any single intervention group, (2) participants receiving the test intervention (i.e., gabapentin or quetiapine), and (3) all trial participants (i.e., frequency in the combined intervention groups); and
- Three different differences between groups, including two in gabapentin sources (no difference in frequency, frequency “higher” in the gabapentin group compared with the placebo group) and two in quetiapine sources (no difference in frequency, frequency “at least twice as high” in the quetiapine group).

We observed 9 combined selection criteria in gabapentin trials and 4 combined selection criteria in quetiapine trials, all in public sources and CSR-synopses (Fig. 1). Each of the selection criteria that we observed included a numerical threshold (albeit different numerical thresholds). Selection criteria did not always include information about participant group or differences between groups; when they did include this information, selection criteria often differed.

When multiple sources described the same trial(s), selection criteria sometimes varied across sources (Additional file 1). Sources about the same trial were sometimes discrepant; in 2/7 (29%) trials with multiple sources that reported selection criteria, we found different sources reported different AEs (Additional file 1). In one additional trial, we found a public source that did not report AEs found in the CSR that should have been reported



in the public source according to the selection criteria described in the public source (Additional file 1).

Number of trials and sources providing AE selection criteria

Although we found no trial with a public source that included all non-systematic AEs (Fig. 2), 47/74 (64%) gabapentin sources and 22/50 (44%) quetiapine sources reported some AE results. Of the sources that did report AE results, 15/47 (32%) gabapentin sources and 19/22 (86%) quetiapine sources reported the selection criteria used.

Selection criteria affect AE reporting

We examined the impact of selection criteria on reporting AEs using all trials for which we had a CSR (Fig. 3). By applying the 45 selection criteria to the AEs reported in eight CSRs, we determined that selection criteria could have a meaningful impact on the number of different AEs that would be reported in other sources (Fig. 3). For example, there were 316 different AEs described in the CSR for Calabrese 2004 [42]. While 126/316 (40%) AEs met the selection criterion “occurring in ≥1% of participants in any treatment group,” only 5/316 (2%) AEs met the selection criterion “occurring in ≥10% of quetiapine-treated patients and twice as frequent in the quetiapine group as the placebo group.” All public sources about Calabrese 2004 that reported selection criteria indicated that they described all AEs occurring in

≥ 10% of quetiapine-treated patients and twice as frequently in the quetiapine group compared with the placebo group.

Discussion

In this study, we examined how reports of clinical trials use selection criteria for reporting non-systematic AEs. We found that public sources and CSR-synopses applied selection criteria while CSRs reported all AEs; consequently, most AEs and serious AEs were not reported in public sources. In public sources, every combined selection criterion we found included a numerical threshold; however, numerical thresholds were not consistent across trials or across sources for individual trials. Some selection criteria also included requirements related to the participant group and the difference between groups.

We found no evidence in public sources (e.g., journal articles) or non-public sources (e.g., CSRs) that any trial used pre-specified selection criteria. We assume that public sources reporting a small number of AEs used selection criteria; however, many public sources did not describe the selection criteria they used. Among public sources that did not describe selection criteria, we could not determine why the authors reported some non-systematic AEs but did not report other non-systematic AEs. Even when selection criteria were described, public sources did not explain why the authors applied those selection criteria (e.g.,

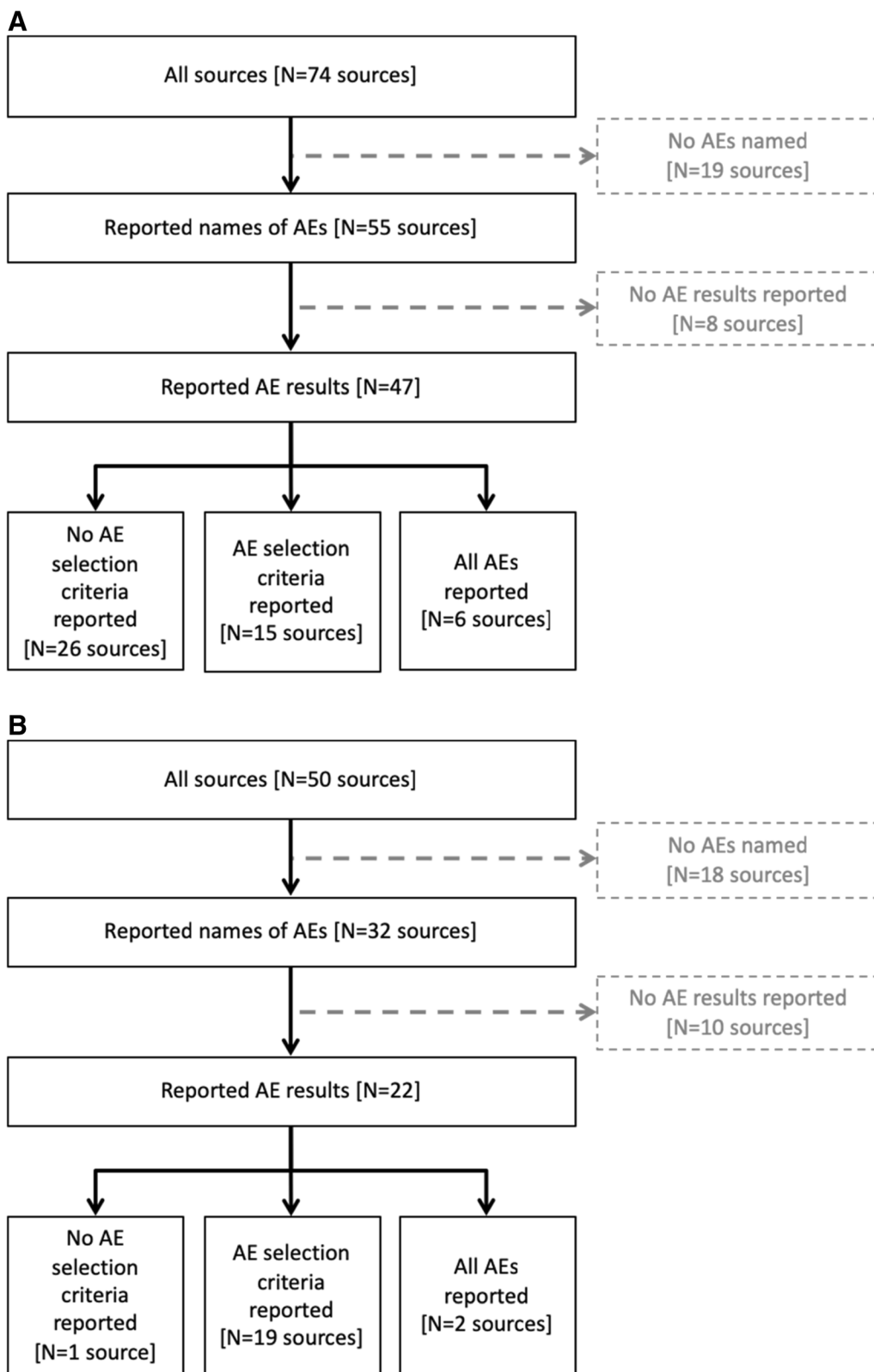
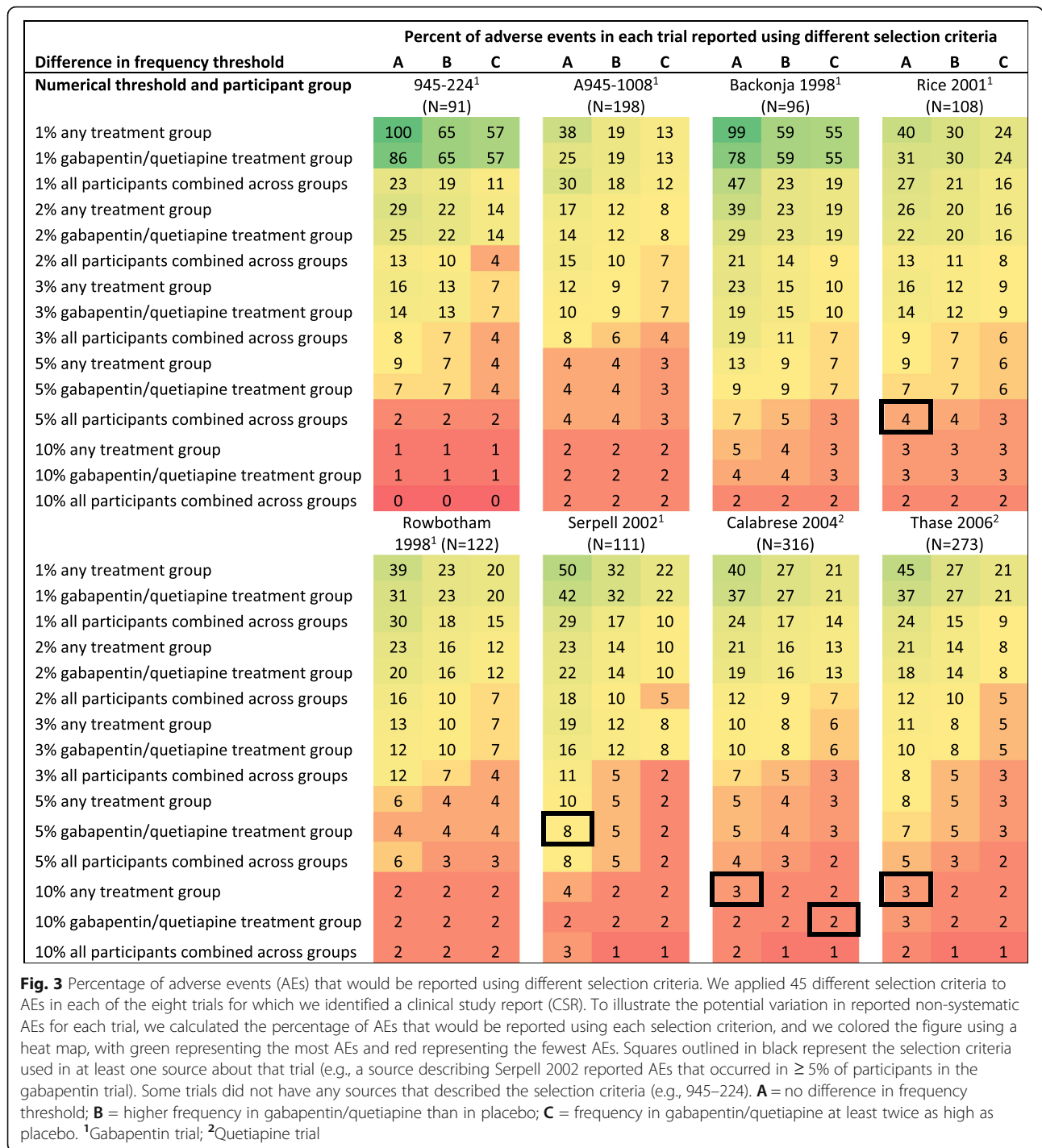


Fig. 2 Reported adverse events (AEs) and adverse event selection criteria. **a** Gabapentin. Six clinical study reports (CSRs) we identified appeared to report all AEs. All other sources in this figure are public sources. **b** Quetiapine. Two CSRs we identified appeared to report all AEs. Two CSR-synopses reported AE results and AE selection criteria. All other sources in this figure are public sources



following a pre-specified statistical analysis plan). Even if investigators could identify which AEs or groups of AEs were most important in a given trial, different selection criteria across trials would make it impossible to ensure unbiased reporting and to conduct unbiased syntheses.

When we applied various selection criteria to eligible trials, we observed meaningful differences in the

number of AEs that would be reported. Trialists might have used selection criteria to identify what they consider the most important AEs; however, we are unaware of any consensus about which AEs are most important for these drugs and conditions. We found no evidence that patients or clinicians were involved in deciding which AEs would be reported in public sources. Instead, it appears that public sources

included the most common AEs. Trialists could have performed analyses similar to ours in which they applied different selection criteria and then decided *post hoc* which selection criteria would allow them to report, or not report, particular AEs. In a previous study, we found that trialists did not “group” AEs for reporting [22], which would be another method to consolidate AEs for reporting that has been advocated elsewhere [37]; however, grouping AEs can also disguise important AEs by combining them with less important AEs [43].

Evidence syntheses (e.g., systematic reviews, clinical practice guidelines) could help identify rare AEs if all observed AEs were available for all trials [44]; however, rare AEs cannot be identified when clinical trials report only those AEs occurring above numerical thresholds. Just as selectively reporting potential benefits based on quantitative results leads to biased meta-analyses [45–48], reporting AEs based on trial results will necessarily lead to biased overall estimates because only “positive” signals will be available. Selection criteria make it impossible for systematic reviewers and meta-analysts to accurately assess the AEs caused by medical interventions.

We found that no public source reported all AEs for any trial. Only serious AEs and AEs occurring in more than 5% of participants are required to be reported in www.ClinicalTrials.gov according to the Food and Drug Administration Amendments Act of 2007 (FDAAA) [1, 49]. Doctors and patients often rely on post-marketing surveillance studies to identify rare AEs, including serious AEs, yet these studies lack comparison groups that are present in clinical trials. It might be possible to detect rare AEs in clinical trials and to calculate between-group differences, if investigators would not use numerical thresholds to determine which AEs to report. Investigators must ensure that for common AEs, such as headache, data are collected consistently across intervention groups so that proportions in each group can be compared. This is difficult when AEs are collected non-systematically. Although systematically collecting AEs would result in more trustworthy and usable information about effects between groups [22, 23], it is not always possible, or even desirable, to anticipate which AEs patients will experience.

Authors have previously discouraged the use of selection criteria [36–38]; however, trials would have to report hundreds or thousands of AEs if selection criteria were not used. Making CSRs and equivalent reports for non-industry trials public, and grouping AEs for analysis and reporting, would partially address the problem of reporting many AEs in journal articles and other public sources. A complete solution to the problems we have identified is not obvious.

Multiple sources of public AE information from trials (e.g., journal articles, FDA reviews), which may be written by different authors for different purposes, lead to inconsistent information for patients and physicians [50]. For example, FDA reviewers have access to CSRs that report all AEs that occurred in each trial. FDA reviewers consider pre-clinical data (e.g., pharmacokinetics, animal trials) and apply clinical and statistical judgment when deciding what to report in medical and statistical reviews about new drugs and biologic agents. The FDA and manufacturers also decide what to include in prescribing information (drug “labels”) written for patients and doctors. By comparison, other stakeholders obtain their information about interventions from a variety of sources, and those sources may use different selection criteria for reporting. For example, individual authors and journal editors can decide what to report in journal articles, which vary tremendously. Different selection criteria across reports of clinical trials, including multiple reports of the same trial, lead to inconsistent and confusing information for stakeholders; consistent standards for reporting AEs, and open access to trial information (e.g., CSRs) could help.

Because current methods for selecting AEs lead to biased reporting that will necessarily produce incorrect effect estimates in systematic reviews and clinical practice guidelines, regulators, lawmakers (e.g., the U.S. Congress) and journals could require that trialists report all AEs on www.ClinicalTrials.gov or other registries. When it is not feasible to report all AEs in a given source (e.g., a conference abstract), the source could direct readers to additional information in a registry such as www.ClinicalTrials.gov. FDA and other regulators could make all AEs publicly available. If all AEs are not made available by regulators or by trialists, systematic reviewers and meta-analysts should interpret results with extreme caution and explain the limitations of using only publicly available data.

There is a pressing need to make clinical trial data available to the public, especially CSRs and IPD; however, sharing CSRs and IPD will not solve all problems identified in our research. First, reports describing hundreds of AEs might overwhelm physicians and patients by including “too much” information [51]. Many AEs reported in CSRs and IPD are not intervention-related; selection criteria might have been used to help decision-makers identify AEs that are caused by medical products. Sharing lengthy and confusing reports and datasets could increase the appearance of transparency while actually disguising important information. Second, reanalyzing clinical trial data is time consuming and therefore expensive, and reanalysis should not be necessary to identify AEs. Most decision-makers want trustworthy *summaries* of clinical trials.

To avoid reporting bias, and to avoid overwhelming patients and physicians with information, trials could pre-specify which AEs will be collected and reported in summaries such as journal articles [36]. Core outcome sets are the minimum sets of outcomes to include in trials [52, 53], and they normally focus on the assessment of potential benefits for people seeing treatment for a particular health problem. Because different types of interventions for the same health problem might be associated with different AEs, and because a single intervention might be used to treat several health problems (e.g., quetiapine is used to treat bipolar depression and schizophrenia), core outcome sets focused on the AEs associated with a particular intervention or group of interventions could also improve the comparability of clinical trials. Notably, core outcome sets for AEs would help trialists identify and report those AEs that are most important to patients, not just the AEs that occurred most often [22, 23, 51]. As described elsewhere, systematic assessment of important AEs would produce much more useful information compared with non-systematic assessment [22, 23].

It is a limitation that we included only two drug-indications, and that we had non-public information for a minority of trials in each case study. The use of selection criteria and the extent to which they are pre-specified and consistent might differ across companies and investigators. Nonetheless, the results of this methodological investigation highlight fundamental problems with methods currently used to report AEs that occur in clinical trials.

Conclusions

Stakeholders require complete AE information to make informed healthcare decisions. Our findings show that systematic reviews and clinical practice guidelines are unlikely to provide accurate effect estimates for AEs because the use of selection criteria results in limited and necessarily biased information about AEs. It may be both difficult and undesirable to include all AEs in journal articles and other public reports, so standards are needed to determine which AEs to include in reports of clinical trials and how to report AEs completely in other public sources such as trial registers [54].

Additional file

Additional file 1: Comparison of adverse events reported in sources about the same trial. (DOCX 266 kb)

Abbreviations

AE: Adverse event; CSR: Clinical study report; FDA: U.S. Food and Drug Administration; FDAAA: Food and Drug Administration Amendments Act of 2007; IPD: Individual patient dataset; MUDS: Multiple Data Sources Study; SRDR: Systematic Review Data Repository

Acknowledgements

We used the open access database, Systematic Review Data Repository (SRDR) and are grateful to Jens Jap, Bryant Smith, and Joseph Lau for making this possible without charge and for the support they provided. The MUDS Investigators include: Lorenzo Bertizzolo, Joseph K. Canner, Terrie Cowley, Kay Dickersin, Peter Doshi, Jeffrey Ehmsen, Nicole Fusco, Gillian Gresham, Nan Guo, Jennifer A. Haythornthwaite, James Heyward, Hwanhee Hong, Tianjing Li, Evan Mayo-Wilson, Jennifer Payne, Diana Pham, Lori Rosman, Elizabeth Stuart, Catalina Suarez-Cuervo, Elizabeth Tolbert, Claire Twose, and S. Swaroop Vedula. Swaroop Vedula and Peter Doshi contributed to the design of the MUDS study and to writing the grant application. Terrie Cowley, Gillian Gresham, James Heyward, Diana Pham, and Elizabeth Tolbert contributed to drafting and finalizing the data extraction forms. Lori Rosman and Claire Taylor designed and ran the electronic searches. Gillian Gresham, James Heyward, Susan Hutless, and Swaroop Vedula screened studies for inclusion. Lorenzo Bertizzolo, Jeffrey Ehmsen, Gillian Gresham, James Heyward, Diana Pham, and Catalina Suarez-Cuervo extracted data.

Authors' contributions

Concept and design: this study was completed as a sub-study of the Multiple Data Sources (MUDS) study, a larger study conceived of and designed as described in the published protocol. EMW drafted the overarching protocol and NF contributed to the protocol and led the design of this sub-study. Acquisition, analysis, or interpretation of data: KD, NF, and TL. EMW contributed to drafting and finalizing the data extraction forms. KD, NF, EMW, and TL screened studies for inclusion. JC, NF, and EMW extracted data. JC, NF, HH, and EMW managed the data and prepared data for analysis. NF and HH analyzed the data. Drafting of the manuscript: NF wrote an initial draft of the paper with EMW and KD. EMW and KD revised the manuscript substantially. Critical revision: all authors reviewed, provided critical revisions, and approved the manuscript for publication.

Funding

This work was supported by contract ME 1303 5785 to Kay Dickersin from the Patient Centered Outcomes Research Institute (PCORI) and funds established for scholarly research on reporting biases at Johns Hopkins by Greene LLP. The funders were not involved in the design or conduct of the study, manuscript preparation, or the decision to submit the manuscript for publication.

Availability of data and materials

The datasets generated and analyzed during the current study and the statistical code are available from the Dryad repository at <https://datadryad.org/resource/doi:10.5061/dryad.mp26fb1>.

Ethics approval and consent to participate

The study received an exemption from the Johns Hopkins Bloomberg School of Public Health Institutional Review Board (IRB No: 00006324).

Consent for publication

Not applicable.

Competing interests

Up to 2008, KD served as an unpaid expert witness for Greene LLP, the plaintiffs' lawyers in litigation against Pfizer that provided several gabapentin documents used for this study. Swaroop Vedula was paid by the plaintiffs' attorneys for research assistance provided to KD for her work as an expert witness. Greene LLP provided a fund for scholarly research on reporting biases that provided salary support for NF and EMW.

Author details

¹Department of Epidemiology and Biostatistics, Indiana University School of Public Health-Bloomington, 1025 E 7th St, #179D, Bloomington, IN 47405, USA. ²Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, 624 N Broadway, Hampton House, Baltimore, MD 21205, USA. ³Department of Surgery, Johns Hopkins School of Medicine, 600 North Wolfe Street, Blalock 1202, Baltimore, MD 21287, USA.

Received: 30 January 2019 Accepted: 16 July 2019

Published online: 05 September 2019

References

- Department of Health and Human Services. 42 CFR Part 11. Clinical trials registration and results information submission: Final rule. 2016. <https://www.federalregister.gov/documents/2016/09/21/2016-22129/clinical-trials-registration-and-results-information-submission>. Accessed 12 May 2018.
- Zarin DA, Tse T, Williams RJ, et al. Trial reporting in ClinicalTrials.gov—the final rule. *N Engl J Med*. 2016;375(20):1998–2004. <https://doi.org/10.1056/NEJMs1611785>.
- Department of Health and Human Services. CFR Title 21, Section 312. Investigational new drug application, 2010.
- European Medicines Agency. Guideline on good pharmacovigilance practices (GVP): Annex I - Definitions (Rev 4), 2017. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/05/WC500143294.pdf. Accessed 12 May 2018.
- Department of Health and Human Services. CFR Title 21, Section 312.32. IND safety reporting. 2010. <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcr/CFRsearch.cfm?CFRPart=312>. Accessed 12 May 2018.
- Golder S, Loke YK, Wright K, et al. Most systematic reviews of adverse effects did not include unpublished data. *J Clin Epidemiol*. 2016;77:125–33. [https://doi.org/10.1016/S0889-8588\(05\)70309-9](https://doi.org/10.1016/S0889-8588(05)70309-9).
- Parsons R, Golder S, Watt I. More than one-third of systematic reviews did not fully report the adverse events outcome. *J Clin Epidemiol*. 2019;108:95–101. <https://doi.org/10.1016/j.jclinepi.2018.12.007>.
- Bonini S, Eichler H-G, Wathion N, et al. Transparency and the European Medicines Agency—sharing of clinical trial data. *N Engl J Med*. 2014;371(26):2452–5.
- Doshi P, Jefferson T. The first 2 years of the European Medicines Agency's policy on access to documents: secret no longer. *JAMA Intern Med*. 2013;173(5):380–2.
- U.S. Food and Drug Administration. FDA Commissioner Scott Gottlieb, M.D., on new steps FDA is taking to enhance transparency of clinical trial information to support innovation and scientific inquiry related to new drugs [press release]. 2018.
- Bierer BE, Li R, Barnes M, et al. A global, neutral platform for sharing trial data. *N Engl J Med*. 2016;374(25):2411–3.
- Krumholz HM, Waldstreicher J. The Yale Open Data Access (YODA) project—a mechanism for data sharing. *N Engl J Med*. 2016;375(5):403–5.
- Ross JS, Waldstreicher J, Bamford S, et al. Overview and experience of the YODA project with clinical trial data sharing after 5 years. *Sci Data*. 2018;5:180268.
- Dickersin K, Min Y-I, Meinert CL. Factors influencing publication of research results: follow-up of applications submitted to two institutional review boards. *JAMA*. 1992;267(3):374–8. <https://doi.org/10.1001/jama.1992.03480030052036>.
- Dwan K, Gamble C, Williamson PR, et al. Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. *PLoS One*. 2013;8(7):e66844. <https://doi.org/10.1371/journal.pone.0066844>.
- Chan AW, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *BMJ*. 2005;330(7494):753–6. <https://doi.org/10.1136/bmj.38356.424606.8F>.
- Chan AW, Hrobjartsson A, Haahr MT, et al. Empirical evidence for selective reporting of outcomes in randomized trials - comparison of protocols to published articles. *JAMA*. 2004;291(20):2457–65. <https://doi.org/10.1001/jama.291.20.2457>.
- Chan AW, Kriebitz-Jeric K, Schmid I, et al. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *Can Med Assoc J*. 2004;171(7):735–40. <https://doi.org/10.1503/cmaj.1041086>.
- Dwan K, Altman DG, Clarke M, et al. Evidence for the selective reporting of analyses and discrepancies in clinical trials: a systematic review of cohort studies of clinical trials. *PLoS Med*. 2014;11(6):e1001666. <https://doi.org/10.1371/journal.pmed.1001666>.
- Mayo-Wilson E, Fusco N, Li T, et al. Multiple outcomes and analyses in clinical trials create challenges for interpretation and research synthesis. *J Clin Epidemiol*. 2017;86:39–50. <https://doi.org/10.1016/j.jclinepi.2017.05.007>.
- Vedula SS, Bero L, Scherer RW, et al. Outcome reporting in industry-sponsored trials of gabapentin for off-label use. *N Engl J Med*. 2009;361(20):1963–71. <https://doi.org/10.1056/NEJMsa0906126>.
- Mayo-Wilson E, Fusco N, Li T, et al. Harms are assessed inconsistently and reported inadequately Part 2: Non-systematic adverse events. *J Clin Epidemiol*. 2019;113:11–9. <https://doi.org/10.1016/j.jclinepi.2019.04.020>.
- Mayo-Wilson E, Fusco N, Li T, et al. Harms are assessed inconsistently and reported inadequately Part 1: Systematic adverse events. *J Clin Epidemiol*. 2019;113:20–7. <https://doi.org/10.1016/j.jclinepi.2019.04.022>.
- Mayo-Wilson E, Li T, Fusco N, et al. Cherry-picking by trialists and meta-analysts can drive conclusions about intervention efficacy. *J Clin Epidemiol*. 2017;91:95–110. <https://doi.org/10.1016/j.jclinepi.2017.07.014>.
- Golder S, Loke YK, Wright K, et al. Reporting of adverse events in published and unpublished studies of health care interventions: a systematic review. *PLoS Med*. 2016;13(9):e1002127. <https://doi.org/10.1371/journal.pmed.1002127>.
- Halfpenny NJA, Quigley JM, Thompson JC, et al. Value and usability of unpublished data sources for systematic reviews and network meta-analyses. *Evid Based Med*. 2016;21(6):208–13. <https://doi.org/10.1136/ebmed-2016-110494>.
- Hart B, Lundh A, Bero L. Effect of reporting bias on meta-analyses of drug trials: reanalysis of meta-analyses. *BMJ*. 2012;344:d7202. <https://doi.org/10.1136/bmj.d7202>.
- Hartung DM, Zarin DA, Guise J-M, et al. Reporting discrepancies between the ClinicalTrials.gov results database and peer-reviewed publications. *Ann Intern Med*. 2014;160(7):477–83. <https://doi.org/10.7326/M13-0480>.
- Jefferson T, Jones MA, Doshi P, et al. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. *Cochrane Database Syst Rev*. 2012;1:CD008965. <https://doi.org/10.1002/14651858.CD008965.pub4>.
- Le Noury J, Nardo JM, Healy D, et al. Restoring Study 329: efficacy and harms of paroxetine and imipramine in treatment of major depression in adolescence. *BMJ*. 2015;351:h4320. <https://doi.org/10.1136/bmj.h4320>.
- Riveros C, Dechartres A, Perrodeau E, et al. Timing and completeness of trial results posted at ClinicalTrials.gov and published in journals. *PLoS Med*. 2013;10(12):e1001566. <https://doi.org/10.1371/journal.pmed.1001566>.
- Rodgers MA, Brown JV, Heirs MK, et al. Reporting of industry funded study outcome data: comparison of confidential and published data on the safety and effectiveness of rhBMP-2 for spinal fusion. *BMJ*. 2013;346:f3981. <https://doi.org/10.1136/bmj.f3981>.
- International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH harmonised tripartite guideline: clinical safety data management: definitions and standards for expedited reporting E2A. 1994. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E2A/Step4/E2A_Guideline.pdf; Accessed 12 May 2018.
- Zarin DA, Tse T, Williams RJ, et al. The ClinicalTrials.gov results database—update and key issues. *N Engl J Med*. 2011;364(9):852–60. <https://doi.org/10.1056/NEJMsa1012065>.
- International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. ICH harmonised tripartite guideline: structure and content of clinical study reports E3. 1995. <http://www.ich.org/products/guidelines/efficacy/efficacy-single/article/structure-and-content-of-clinical-study-reports.html>. Accessed 12 May 2018.
- Lineberry N, Berlin JA, Mansi B, et al. Recommendations to improve adverse event reporting in clinical trial publications: a joint pharmaceutical industry/journal editor perspective. *BMJ*. 2016;355:i5078. <https://doi.org/10.1136/bmj.i5078>.
- Honvo G, Bannuru RR, Bruyère O, et al. Recommendations for the reporting of harms in manuscripts on clinical trials assessing osteoarthritis drugs: a consensus statement from the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO). *Drugs Aging*. 2019;36(1):145–59.
- Ioannidis JP, Evans SJ, Gøtzsche PC, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med*. 2004;141(10):781–8. <https://doi.org/10.7326/0003-4819-141-10-200411160-00009>.
- Mayo-Wilson E, Hutfless S, Li T, et al. Integrating multiple data sources (MUDS) for meta-analysis to improve patient-centered outcomes research: a protocol for a systematic review. *Syst Rev*. 2015;4:143. <https://doi.org/10.1186/s13643-015-0134-z> (Correction: 2018; 7:48).
- Mayo-Wilson E, Doshi P, Dickersin K. Are manufacturers sharing data as promised? *BMJ*. 2015;351:h4169. <https://doi.org/10.1136/bmj.h4169>.
- Stata Statistical Software: Release 14 [program]. College Station; 2015.
- AstraZeneca. A multicenter, double-blind, randomized, placebo-controlled, double-dummy trial of the use of quetiapine fumarate (SEROQUEL

- [Registered]) in the treatment of patients with bipolar depression. Study code: 5077US/0049, 2005.
43. Crowe BJ, Xia HA, Berlin JA, et al. Recommendations for safety planning, data collection, evaluation and reporting during drug, biologic and vaccine development: a report of the safety planning, evaluation, and reporting team. *Clin Trials*. 2009;6(5):430–40.
 44. Chalmers I. Proposal to outlaw the term “negative trial”. *BMJ*. 1985; 290(6473):1002.
 45. Hutton J, Williamson PR. Bias in meta-analysis due to outcome variable selection within studies. *J Roy Stat Soc C App*. 2000;49(3):359–70. <https://doi.org/10.1111/1467-9876.00197>.
 46. Kirkham JJ, Dwan KM, Altman DG, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ*. 2010;340:c365. <https://doi.org/10.1136/bmj.c365>.
 47. Williamson PR, Gamble C. Identification and impact of outcome selection bias in meta-analysis. *Stat Med*. 2005;24(10):1547–61. <https://doi.org/10.1002/sim.2025>.
 48. Williamson PR, Gamble C, Altman DG, et al. Outcome selection bias in meta-analysis. *Stat Methods Med Res*. 2005;14(5):515–24. <https://doi.org/10.1191/0962280205sm415oa> [published Online First: 2005/10/27].
 49. 110th Congress of the United States of America. Food and Drug Administration Amendments Act, Section 801. 2007.
 50. Mayo-Wilson E, Li T, Fusco N, et al. Practical guidance for using multiple data sources in systematic reviews and meta-analyses (with examples from the MUDS study). *Res Synth Methods*. 2018;9(1):2–12. <https://doi.org/10.1002/jrsm.1277>.
 51. Mayo-Wilson E, Golozar A, Cowley T, et al. Methods to identify and prioritize patient-centered outcomes for use in comparative effectiveness research. *Pilot Feasibility Stud*. 2018;4(95). <https://doi.org/10.1186/s40814-018-0284-6>.
 52. Prinsen CA, Vohra S, Rose MR, et al. Core Outcome Measures in Effectiveness Trials (COMET) initiative: protocol for an international Delphi study to achieve consensus on how to select outcome measurement instruments for outcomes included in a ‘core outcome set’. *Trials*. 2014; 15(1):247.
 53. COMET Initiative. <http://www.comet-initiative.org/>. Accessed 12 May 2018.
 54. Dickersin K, Mayo-Wilson E. Standards for design and measurement would make clinical research reproducible and usable. *P Natl Acad Sci USA*. 2018; 115(11):2590–4. <https://doi.org/10.1073/pnas.1708273114>.

Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

