Journal of Clinical Epidemiology 113 (2019) 20–27

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Harms are assessed inconsistently and reported inadequately part 1: systematic adverse events

Evan Mayo-Wilson[a,*,1], Nicole Fusco[a,1], Tianjing Li[a], Hwanhee Hong[b], Joseph K. Canner[c], Kay Dickersin[a], for the MUDS investigators

[a]*Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, MD 21205, USA*
[b]*Department of Mental Health, Johns Hopkins University Bloomberg School of Public Health, 624 N Broadway, Hampton House, Baltimore, MD 21205, USA*
[c]*Department of Surgery, Johns Hopkins University School of Medicine, 600 North Wolfe Street, Blalock 1202, Baltimore, MD 21287, USA*

Accepted 5 April 2019; Published online 2 May 2019

Abstract

Objectives: We examined systematic adverse events (AEs) in Part 1 (of 2) of a study describing the assessment and reporting of AEs in clinical trials.

Study Design and Setting: We examined 52 public and nonpublic data sources about trials of quetiapine for bipolar depression using data from the Multiple Data Sources study. We extracted and compared information about systematic AEs (i.e., AEs assessed for all participants) in six prespecified domains: cardiovascular, cholesterol, endocrine, extrapyramidal symptoms, mania, and weight.

Results: Eligible trials did not assess and report the same systematic AEs, and most results were not available in public sources. Overall, public sources reported 159 results, of which 92 of 159 (58%) included sufficient statistical information to calculate the treatment effect and its precision. Nonpublic sources reported 636 results; 630 of 636 (99%) reported sufficient statistical information.

Conclusion: Systematic AEs were defined and analyzed in many ways, which led to many numerical results. Most systematic AEs were not mentioned in public sources. To minimize bias, methods for defining and analyzing potential AEs should be prespecified in trial registers and protocols. All trial results should be publicly available so that stakeholders can compare benefits and AEs. Trials could report core sets of AEs to facilitate decision-making.

Keywords: Adverse events; Clinical trials; Drug safety; Reporting bias; Open science; Harms

Funding: This study was supported by contract ME 1303 5785 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.

Up to 2008, K.D. served as an unpaid expert witness for 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 Kay Dickersin for her work as an expert witness. Greene LLP, a law firm that has represented various clients in litigation against Pfizer, including the case in which Kay Dickersin served as an expert witness, provided a fund for scholarly research on reporting biases that partially covered salaries of some investigators performing analysis and interpretation of data. For the collaborators: Terrie Cowley is the president of The TMJ Association, Ltd., which received funding from Pfizer from 2009 to 2011 (for the Campaign to End Chronic Pain in Women) and from Purdue Pharma, L.P. in 2008, 2010, and 2012 to 2016 (to support the TMJ Association science meeting and develop educational materials based on scientific research advances). Peter Doshi was a corecipient of a UK National Institute for Health Research grant (HTA—10/80/01 Update and amalgamation of two Cochrane Reviews: neuraminidase inhibitors for preventing and treating influenza in healthy adults and children—http://www.nets.nihr.ac.uk/projects/hta/108001). Peter Doshi received €1,500 from the European Respiratory Society in support of his travel to the society's September 2012 annual congress in Vienna, where he gave an invited talk on oseltamivir. Peter Doshi gratefully acknowledges the American Association of Colleges of Pharmacy for its funding support ($11,000) for a study to analyze written medical information regarding the possible harms of statins. AACP had no involvement in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of this manuscript. Peter Doshi is also an associate editor of The BMJ and an unpaid member of the IMEDS steering committee at the Reagan-Udall Foundation for the FDA, which focuses on drug safety research.

Data sharing: Statistical code and datasets are available from the Dryad repository, https://doi.org/10.5061/dryad.mp26fb1.

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

[1] These authors contributed equally to this article.

* Corresponding author. Tel.: +1 443 287 5042.
E-mail address: evan.mayo-wilson@jhu.edu (E. Mayo-Wilson).

https://doi.org/10.1016/j.jclinepi.2019.04.022

What is new?

- Adverse events (AEs) can be assessed systematically in clinical trials by collecting information for all participants using clinical examinations, questionnaires, and medical instruments.
- Systematic AEs can be defined using the five elements of an outcome: (1) domain, (2) specific measurement, (3) specific metric, (4) method of aggregation, and (5) time-point.
- Multiple outcome definitions and multiple methods of analysis create opportunities to "cherry pick" systematic AE results, consistent with previous studies about potential benefits of treatments.
- To prevent reporting bias, systematic AEs should be prespecified in trial registrations, protocols, and statistical analysis plans. All adverse events that occur in clinical trials should be reported publicly (e.g., on www.ClinicalTrials.gov).
- Core outcome sets should identify AEs that are important to patients and recommend that those AEs be assessed systematically.

1. Introduction

To make informed decisions about health interventions, patients and other stakeholders need accurate and complete information about potential benefits and adverse events (AEs). Investigators assess potential benefits systematically in clinical trials; that is, investigators assess potential benefits for all participants using standardized approaches such as clinical examinations, questionnaires, and medical instruments. Likewise, investigators may assess AEs systematically, especially when AEs are anticipated because of preclinical studies or because of evidence from related clinical trials (Table 1) [1]. By contrast, nonsystematic AEs are not assessed for every participant; instead, participants report nonsystematic AEs in response to questions like "Have you noticed any symptoms since your last examination?" [2]. Previous studies have described nonsystematic AEs as "passively collected" and systematic AEs as "proactively collected" [3].

Previous studies show that potential benefits can be reported in ways that are misleading. That is, multiple outcome definitions and multiple methods of analysis allow investigators to "cherry pick" results [4–6]. If systematic AEs are defined and analyzed similar to potential benefits, then cherry-picking results for systematic AEs could lead to incorrect estimates of their likelihood and severity.

Part 1 and Part 2 of this study about AEs are substudies of the Multiple Data Sources (MUDS) study, which followed a published protocol [7] and protocol amendments [4]. Our objectives were to (1) describe the assessment and reporting of systematic AEs in eligible clinical trials and (2) compare reporting of AEs in public and nonpublic sources. We assessed systematic AEs (Part 1) apart from nonsystematic AEs (Part 2) because differences in data collection and analysis might contribute to differences in reporting.

2. Methods

2.1. Eligible trials and sources

Eligible studies were parallel, randomized clinical trials that evaluated either gabapentin for neuropathic pain or quetiapine for bipolar depression, compared with placebo. We selected these case studies because we had access to both public and nonpublic sources for some of the eligible trials. We searched for public and nonpublic sources (Table 1) and requested additional information from manufacturers of gabapentin and quetiapine [15]. Here, we describe the seven quetiapine trials (52 sources); no gabapentin studies were included in this substudy because gabapentin trials did not report any systematic AEs. The 52 eligible sources for quetiapine trials include individual patient data (IPD) for two trials; we found systematic AE data in one dataset but not in the other dataset; therefore, 51 sources (including one IPD) were analyzed for this substudy.

2.2. Data extraction

From each source of each eligible trial, two investigators independently extracted data using the open access Systematic Review Data Repository (http://srdr.ahrq.gov/) and resolved differences by discussion. 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 "nonsystematic" (Table 1). For the time point closest to 8 weeks, we extracted all systematic AE definitions and results (Table 1) for six prespecified AE domains that patient and clinician coinvestigators confirmed are important to people with bipolar disorder: cardiovascular, cholesterol, endocrine, extrapyramidal symptoms, mania, and weight.

Just as investigators define potential benefits using five elements [4,6,13,16,17], we defined systematic AEs using five elements: (1) domain (e.g., mania); (2) specific measurement (e.g., Young Mania Rating Scale); (3) metric (e.g., change from baseline); (4) method of data aggregation (e.g., mean); and (5) time point at which the AE was assessed.

We extracted results that were reported separately for each trial; we did not extract results of analyses that pooled multiple trials.

For each reported result, we assessed whether we could estimate the size of the difference between intervention arms and the precision of that estimate (Table 1).

Table 1. Glossary of terms

Term	Definition/example
Terms related to adverse events (AEs)	
Adverse event (AE)[1]	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'' [8]. The U.S. Food and Drug Administration (FDA) and other regulators use this definition [9,10].
Nonsystematic adverse events	According to The Final Rule [11,12], '''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 like ''Have you noticed any symptoms since your last examination?''
Non-unique AE	In the Multiple Data Sources (MUDS) study, a ''non-unique'' AE is an AE counted for each source in which it appears.
Result	In the MUDS study, a ''result'' is a numerical contrast between a treatment and comparison arm (e.g., relative risk, mean difference).
Systematic adverse events	According to The Final Rule [11,12], '''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 [13]. For example, ''proportion of participants with 50% change from baseline to 8 weeks on the Young Mania Rating Scale total score.''
Unique AE	In the MUDS study, a ''unique'' AE is a defined AE counted only once, regardless of how many times it appeared in all sources.
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 [14].
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 participant data (IPD)	A table or database in which each record contains data for a single participant [14].
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.

2.3. Reanalysis of IPD

For continuous AEs included in IPD, we calculated the mean change from baseline. For dichotomous AEs, we calculated the proportion of participants who experienced the AE. We replicated the methods for handling missing data that were described in corresponding CSRs. We performed all analyses in Stata 14 [18].

2.4. Comparison of reporting across multiple sources

We compared the following information in public and nonpublic sources:

Table 2. Trials reported different adverse event domains

	Prespecified AE domains					
Trial identifier	**Cardiovascular**	**Cholesterol**	**Endocrine**	**Extrapyramidal symptoms**	**Mania**	**Weight**
Calabrese, 2004				✓✓	✓✓	✓✓
Gao, 2014				✓	✓✓	
Li, 2014					✓	
McElroy 2010		✓✓	✓✓	✓✓	✓	✓✓
Suppes, 2010		✓✓		✓	✓	✓✓
Thase, 2006				✓✓	✓✓	✓✓
Young, 2008		✓✓	✓✓	✓	✓	✓✓

✓ represents at least one public source reported at least one result.
✓✓ represents at least one public source reported at least one result for which we could calculate the between-arm effect and its precision.

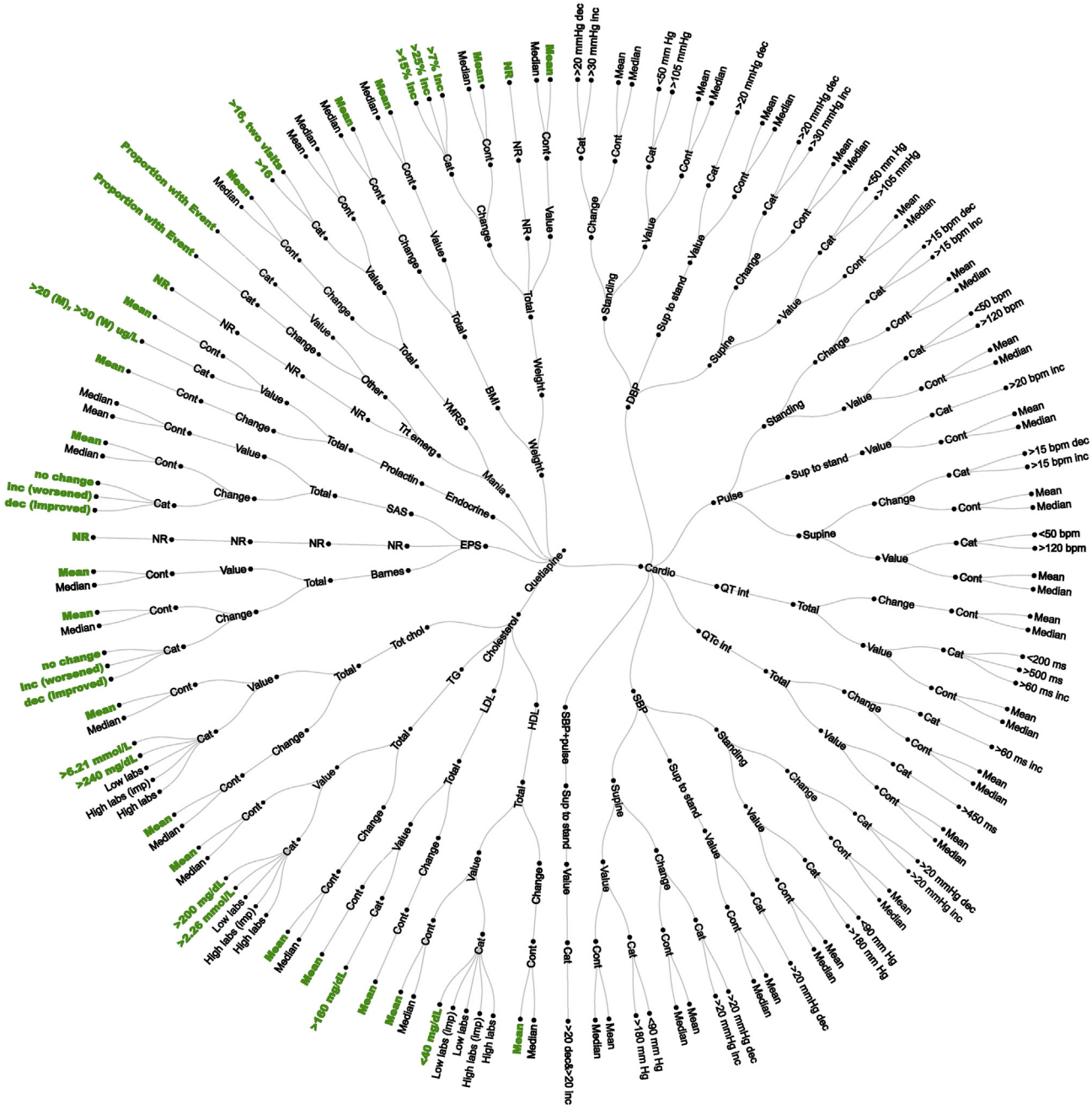

Fig. 1. Multiple outcome definitions for systematic adverse events. Outcomes available in public sources only or in both public and nonpublic sources are in green, whereas outcomes available only in nonpublic sources are in black. Although the total number of unique outcomes was not sensitive to the way we classified domains and measures, the number of domains and measures were related to the way we classified measures. For example, we treated LDL and HDL as separate measures; if we had treated LDL and HDL as subscales of total cholesterol, we would find different numbers of measures and subscales but the same number of unique outcomes. Barnes, Barnes Akathisia Rating Scale; BMI, body mass index; Cardio, cardiovascular; Cat, categorical; Change, change from baseline; Cont, continuous; DBP, diastolic blood pressure; dec, decrease; EPS, extrapyramidal symptoms; HDL, high-density lipoprotein; High labs, elevated laboratory values; High labs (imp), elevated laboratory values, but improved since beginning of study; inc, increase; LDL, low-density lipoprotein; Low labs, laboratory values below normal range; Low labs (imp), laboratory values below normal range, but improved since beginning of study; (M), men; NR, not reported; QT int, QT interval; QTc int, QTc interval; SAS, Simpson-Angus Scale; SBP, systolic blood pressure; Sup to stand, moving from supine to standing; Tot chol, total cholesterol; Trt emerg, treatment emergent; TG, triglycerides; Value, value at a time point; (W), women; YMRS, Young's Mania Rating Scale. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

(1) The number of defined systematic AEs (i.e., an AE was defined if all five elements of the outcome were reported, Table 1);

(2) The number of unique systematic AEs (i.e., we counted "unique" systematic AEs once regardless of how many times they appeared across sources, Table 1);

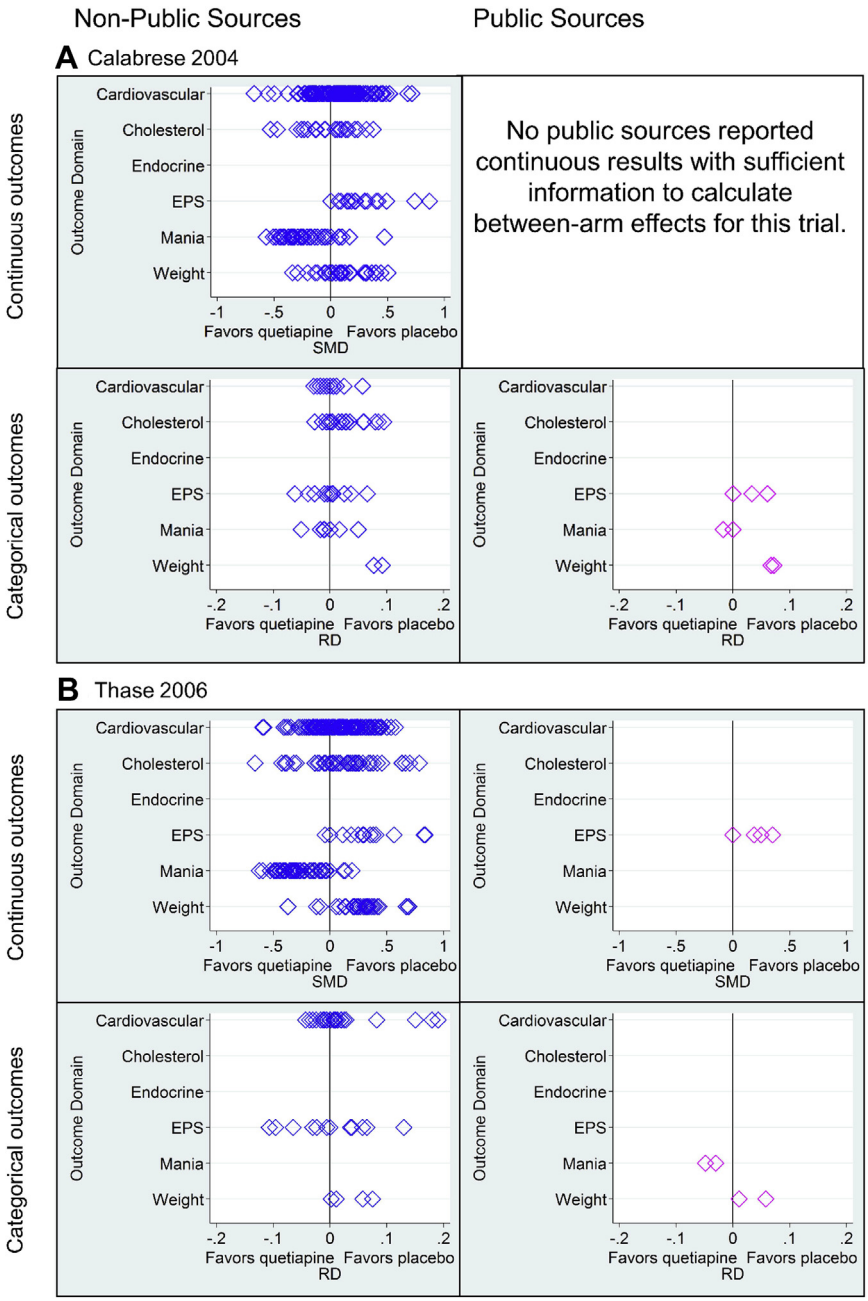

Fig. 2. Systematic adverse events in public and nonpublic sources for two trials: (A) Calabrese 2004. (B) Thase 2006. This figure includes the two trials that reported meta-analyzable results in both public and nonpublic sources. Nonpublic sources about each trial included hundreds of results; by contrast, public sources included fewer than 10 results and omitted AE domains that were included in nonpublic sources. EPS, extrapyramidal symptoms; RD, risk difference (used for dichotomous outcomes); SMD, standardized mean difference (used for continuous outcomes).

(3) The number of nonunique systematic AEs (i.e., systematic AEs that we counted each time they appeared, Table 1); and
(4) The proportion of results for which we could calculate the between-arm effect (e.g., mean difference between arms, risk difference) and its precision (Table 1).

Finally, we plotted the results we extracted from public and nonpublic sources. For continuous AEs, we plotted the standardized mean difference. For dichotomous AEs, we plotted the risk difference and conducted a sensitivity analysis plotting the risk ratio.

3. Results

3.1. Trials of the same drug and condition did not assess the same AEs

Eligible trials did not assess the same systematic AEs. We were able to determine which AEs were assessed systematically in two trials for which we had CSRs. For the

Table 3. More outcomes and results were reported in nonpublic sources compared with public sources

	Public sources *n* (%)	Nonpublic sources *n* (%)	All sources *n*
Number of sources	46	6	52[a]
Number of outcomes			
Prespecified AE domains	5 (83)	5 (83)	6
Unique AEs	38 (30)	113 (89)	127
Number of results			
Nonunique results	159 (20)	636 (80)	795

Abbreviation: AE, adverse events.

[a] The 52 sources include individual participant data for two trials, of which only one included systematic AEs.

Public sources: Journal articles, conference abstracts, commentaries, posters, trial registrations and associated results, and medical and statistical reviews created by the Food and Drug Administration; Nonpublic sources: Individual patient data, clinical study reports, and clinical study report-synopses.

five trials for which we could not obtain CSRs, we were unable to differentiate between (1) systematic AEs that were not assessed and (2) systematic AEs that were assessed but not reported.

All seven trials assessed mania; however, only three of seven (43%) trials publicly reported results for which we could calculate the between-arm effect and its precision (Table 2). Six (86%) trials publicly reported extrapyramidal symptoms; because we did not have nonpublic sources, it was unclear whether the seventh trial assessed extrapyramidal symptoms. For the weight and cholesterol domains, five of seven (71%) trials reported at least one result for which we could calculate the between-arm effect and its precision; however, only three of seven (43%) trials publicly reported effects on cholesterol. The two trials for which we had CSRs did *not* assess endocrine effects, but two other trials assessed this domain. The two trials for which we had CSRs assessed cardiovascular effects, but no trial publicly reported cardiovascular effects.

3.2. Public sources reported AEs inadequately

Within each domain, trials used multiple specific measurements, metrics, and methods of aggregation. For example, the domain "extrapyramidal symptoms" was measured using both the Barnes Akathisia Rating Scale and the Simpson-Angus Scale (Fig. 1, Appendix 1).

In addition, multiple results were sometimes associated with the same AE because of variations in the methods of analysis [6]. Within each domain, results in nonpublic sources varied in both magnitude and direction; that is, some results favored quetiapine, and some results favored placebo (Fig. 2). We found few results in public sources, and there was little variation among them (Fig. 2). Findings from Figure 2 were substantially unchanged when we analyzed the data using the risk ratio rather than the risk difference (Appendix 2).

We found 127 unique systematic AEs in all sources, of which 38 of 127 (30%) were reported in public sources, and 113 of 127 (89%) were reported in nonpublic sources (Table 3).

Compared with results in nonpublic sources, results in public sources were incomplete. Of the nonunique results in public and nonpublic sources, 92 of 159 (58%) and 630 of 636 (99%) included sufficient information to calculate the between-arm effect and its precision, respectively.

4. Discussion

In Part 1 of this study, we found that trials of the same intervention for the same health problem did not collect and report the same systematic AEs. Moreover, investigators could make a drug that actually increases risk of harm appear to reduce risk of harm by changing the outcome definition. The ability to manipulate results within prespecified outcome domains is concerning because most information about systematic AEs is not available to decision-makers. Although our analysis focused on a single intervention, our findings are consistent with previous studies about AE assessment and reporting [4,6,19–21]. Furthermore, multiplicity of systematic AE definitions and analyses, and the underreporting of associated results, mirror the

Box 1 Recommendations for adverse event assessment and reporting

1. Use preclinical studies, trials of other drugs in the same class, and observational studies to anticipate adverse events (AEs) that might occur in a clinical trial;
2. Collect data on anticipated AEs systematically in clinical trials;
3. For each systematically collected AE, prespecify in protocols and statistical analysis plans all five elements of the outcome, the methods of analysis, and the methods for data collection;
4. Develop and use core outcome sets for AEs for interventions or classes of interventions (e.g., drug class);
5. Make all results publicly available.

ways in which potential benefits are assessed and reported [4,6]. Our study identified several ways to improve the collection and reporting of AEs in clinical trials (Box 1).

AEs can be assessed systematically in clinical trials. Although many AEs are assessed nonsystematically as described in Part 2 of this study [2], anticipated AEs can be assessed by recording information for all participants using prespecified methods. Systematically assessing AEs that matter to patients could improve the accuracy of estimates within trials, and systematically assessing the severity and duration of AEs would be more informative than nonsystematically recording the mere occurrence of events [22]. For example, a patient who wants to use the antipsychotic associated with the least weight gain needs systematic information about weight gain in clinical trials, not merely evidence that several antipsychotics are associated with weight gain. Just as clinical trials provide trustworthy information about which interventions are most effective for reducing symptoms of disease (e.g., reducing depression), assessing AEs systematically would provide trustworthy evidence about which interventions cause the most harm. Systematic collection might even prevent nonserious AEs from becoming serious; for example, better monitoring the effects of interventions that cause dizziness or increase blood glucose could help patients and clinicians discontinue those interventions before they cause injurious falls and diabetes. Systematic AEs could be selected based on existing evidence about AEs that are most likely to occur, and they could be selected based on patient preferences [2]. Over time, trials might include more or different systematic AEs as investigators learn about the effects of interventions.

Systematic AEs can be defined using five elements, which should be prespecified [13,16]. We found that results in a trial for a single AE domain could differ in both magnitude and direction, which is consistent with previous evidence about potential benefits [6]. If the five elements of each outcome and the methods of analysis are not prespecified in protocols and statistical analysis plans, then trialists and systematic reviewers can ‘‘cherry-pick’’ systematic AEs based on their results [5,6,23].

Systematic assessment of AEs would improve systematic reviews and clinical guidelines. It may be inappropriate to synthesize nonsystematic AEs across trials (e.g., using meta-analysis) because nonsystematic AEs are reported selectively depending on post-hoc selection criteria [2]. Furthermore, systematic and nonsystematic methods to assess AEs can lead to different results [24–26]; therefore, it may be inappropriate to synthesize AEs that have been defined, collected, and analyzed using different methods. Information about AEs will be most useful if AEs are assessed systematically and consistently across trials of the same interventions and conditions [22].

Including systematic AEs in core outcome sets, which are the minimum sets of outcomes to include in trials of interventions or conditions [27,28], could improve consistency across trials and improve healthcare decisions. Core outcome sets for particular diseases or conditions often focus on potential benefits, yet the same principles used to identify core benefits could be extended to identify core harms [22]. Because different types of interventions for the same condition might cause different AEs (e.g., depression might be treated using psychological or pharmacologic interventions, which are associated with qualitatively different harms) and because a single intervention can be used to treat several conditions (e.g., quetiapine is used to treat bipolar depression and schizophrenia), it might be appropriate to have core sets of AEs for interventions or classes of interventions (e.g., drug classes) rather than core sets of AEs for health conditions.

Information about the collection of AEs and the results of all AE analyses should be publicly accessible. The Consolidated Standards of Reporting Trials extension for harms provides some reporting standards [29], and other reporting guidelines have suggested that AEs be prespecified [3]. When clinical trials do not assess the same outcomes or do not report their results completely, systematic reviewers and health care decision-makers must rely on subsets of the available trials, the results of which might not represent the true benefits and harms of interventions [5,30].

CRediT authorship contribution statement

Evan Mayo-Wilson: Conceptualization, Data curation, Investigation, Methodology, Project administration, Writing - original draft, Writing - review & editing. **Nicole Fusco:** Conceptualization, Data curation, Formal analysis, Investigation, Visualization, Methodology, Writing - original draft, Writing - review & editing. **Tianjing Li:** Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Writing - original draft, Writing - review & editing. **Hwanhee Hong:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing - review & editing. **Joseph K. Canner:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing - review & editing. **Kay Dickersin:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Writing - review & editing.

Acknowledgments

The authors 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 authors acknowledge and thank the MUDS Investigators. Swaroop Vedula and Peter Doshi contributed to the initial design of the MUDS study and to writing the grant application. Terrie Cowley, Gillian Gresham, Jennifer A. Haythornthwaite, James Heyward, Diana Pham, Jennifer

Payne, 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 Hutfless, and Swaroop Vedula screened studies for inclusion. Lorenzo Bertizzolo, Jeffrey Ehmsen, Gillian Gresham, James Heyward, Diana Pham, and Catalina Suarez-Cuervo extracted data.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2019.04.022.

References

[1] Department of Health and Human Services. CFR Title 21, Section 312. Investigational New Drug Application; 2010.

[2] Mayo-Wilson E, Fusco N, Li T, Hong H, Canner JK, Dickersin K, et al. Harms are assessed inconsistently and reported inadequately Part 2: non-systematic adverse events. J Clin Epidemiol 2019. Submitted.

[3] Lineberry N, Berlin JA, Mansi B, Glasser S, Berkwits M, Klem C, et al. Recommendations to improve adverse event reporting in clinical trial publications: a joint pharmaceutical industry/journal editor perspective. BMJ 2016;355:i5078.

[4] Mayo-Wilson E, Li T, Fusco N, Bertizzolo L, Canner JK, Cowley T, et al. Cherry-picking by trialists and meta-analysts can drive conclusions about intervention efficacy. J Clin Epidemiol 2017;91:95−110.

[5] Li T, Mayo-Wilson E, Fusco N, Hong H, Dickersin K. Caveat emptor: the combined effects of multiplicity and selective reporting. Trials 2018;19:497.

[6] Mayo-Wilson E, Fusco N, Li T, Hong H, Canner J, Dickersin K. Multiple outcomes and analyses in clinical trials create challenges for interpretation and research synthesis. J Clin Epidemiol 2017;86:39−50.

[7] Mayo-Wilson E, Hutfless S, Li T, Gresham G, Fusco N, Ehmsen J, 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. Correction: 2018; 7:48.

[8] International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use. In: ICH harmonised tripartite guideline: clinical safety data management: definitions and standards for expedited reporting E2A. Available at http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E2A/Step4/E2A_Guideline.pdf. Accessed May 12, 2018.

[9] European Medicines Agency. Guideline on good pharmacovigilance practices (GVP): Annex I - Definitions (Rev 4). 2017. Available at http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/05/WC500143294.pdf. Accessed May 12, 2018.

[10] Department of Health and Human Services. CFR Title 21, Section 312.32. IND safety reporting 2010. Available at https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRsearch.cfm?CFRPart=312. Accessed May 12, 2018.

[11] Department of Health and Human Services. 42 CFR Part 11. Clinical trials registration and results information submission: final rule. 2016. Available at https://www.federalregister.gov/documents/2016/09/21/2016-22129/clinical-trials-registration-and-results-information-submission. Accessed May 12, 2018.

[12] Zarin DA, Tse T, Williams RJ, Carr S. Trial reporting in ClinicalTrials. gov—the final rule. N Engl J Med 2016;375:1998−2004.

[13] Zarin DA, Tse T, Williams RJ, Califf RM, Ide NC. The ClinicalTrials.gov results database–update and key issues. N Engl J Med 2011;364:852−60.

[14] International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use. In: ICH harmonised tripartite guideline: structure and content of clinical study reports E3. Available at http://www.ich.org/products/guidelines/efficacy/efficacy-single/article/structure-and-content-of-clinical-study-reports.html. Accessed May 12, 2018.

[15] Mayo-Wilson E, Doshi P, Dickersin K. Are manufacturers sharing data as promised? BMJ 2015;351:h4169.

[16] Cybulski L, Mayo-Wilson E, Grant S. Improving transparency and reproducibility through registration: the status of intervention trials published in clinical psychology journals. J Consult Clin Psychol 2016;84(9):753.

[17] Saldanha IJ, Dickersin K, Wang X, Li T. Outcomes in cochrane systematic reviews addressing four common eye conditions: an evaluation of completeness and comparability. PLoS One 2014;9:e109400.

[18] StataCorp LP. Stata Stat Softw Release 14 2015: College Station, TX.

[19] Doshi P, Jefferson T, Del Mar C. The imperative to share clinical study reports: recommendations from the Tamiflu experience. PLoS Med 2012;9(4):e1001201.

[20] Golder S, Loke YK, Bland M. Unpublished data can be of value in systematic reviews of adverse effects: methodological overview. J Clin Epidemiol 2010;63:1071−81.

[21] Golder S, Loke YK, Wright K, Norman G. Reporting of adverse events in published and unpublished studies of health care interventions: a systematic review. PLoS Med 2016;13(9):e1002127.

[22] Mayo-Wilson E, Golozar A, Cowley T, Fusco N, Gresham G, Haythornthwaite J, et al. Methods to identify and prioritize patient-centered outcomes for use in comparative effectiveness research. Pilot Feasibility Stud 2018;4:95.

[23] Mayo-Wilson E, Li T, Fusco N, Dickersin K, investigators M. 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.

[24] Allen EN, Barnes KI, Mushi A, Massawe I, Staedke SG, Mehta U, et al. Eliciting harms data from trial participants: how perceptions of illness and treatment mediate recognition of relevant information to report. Trials 2011;12(Suppl 1):A10.

[25] Bent S, Padula A, Avins AL. Brief communication: better ways to question patients about adverse medical events: a randomized, controlled trial. Ann Intern Med 2006;144:257−61.

[26] Safer DJ. Design and reporting modifications in industry-sponsored comparative psychopharmacology trials. J Nerv Ment Dis 2002;190(9):583−92.

[27] COMET initiative. Available at http://www.comet-initiative.org/. Accessed May 12, 2018.

[28] Prinsen CA, Vohra S, Rose MR, King-Jones S, Ishaque S, Bhaloo Z, 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.

[29] Ioannidis JP, Evans SJ, Gøtzsche PC, O'neill RT, Altman DG, Schulz K, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Ann Intern Med 2004;141:781−8.

[30] Dickersin K, Mayo-Wilson E. Standards for design and measurement would make clinical research reproducible and useable. Proc Natl Acad Sci U S A 2018;115(11):2590−4.