Journal of Clinical Epidemiology 113 (2019) 11−19

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Harms are assessed inconsistently and reported inadequately Part 2: nonsystematic 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

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

Study Design and Setting: We examined 21 trials of gabapentin for neuropathic pain (52 sources) and seven trials of quetiapine for bipolar depression (80 sources) using data from the Multiple Data Sources study. We extracted and compared information about nonsystematic AEs (i.e., AEs that were not assessed for every participant), including AEs categorized as ''serious.'' We recorded whether AEs were grouped by anatomic or physiological system.

Results: Trials of the same drug reported information about different AEs. Information in public sources was inadequate for decision-making. No public source reported all AEs, or all serious AEs, identified in nonpublic sources about the same trial. Of trials with only public sources, 2/15 (13%) gabapentin and 0/3 (0%) quetiapine trials grouped AEs by anatomic or physiological system.

Conclusion: Public sources contained little information about nonsystematic AEs, including serious AEs. Grouping might make nonsystematic AEs easier to detect; however, most public sources did not report grouped AEs. Standards are needed to improve the collection and reporting of nonsystematic AEs so that stakeholders can use trials to assess the balance of potential benefits and harms.

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

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

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 University by Greene LLP. The funders were not involved in the design or conduct of the study, article preparation, or the decision to submit the article for publication.

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

Conflict of interest: 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 K.D. 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 K.D. 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 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 article. 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.

[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.2019.04.020

What is new?

- The presence or absence of a nonsystematic adverse event (AEs) is not assessed for all participants in a clinical trial; instead, nonsystematic AEs are reported to investigators in response to open-ended questions such as ‘‘have you noticed any symptoms since your last visit?’’
- Trials of the same intervention for the same condition include information about different nonsystematic AEs.
- Most nonsystematic AEs, including serious AEs, were not reported in public sources; this suggests that nonsystematic AEs might be ‘‘cherry-picked’’ for reporting.
- Nonsystematic AEs can be grouped for analysis (e.g., to increase statistical power), but few public sources reported grouped AEs.
- Nonsystematic assessment leads to numerous problems for interpretation and synthesis. All AEs that occur in clinical trials should be reported publicly (e.g., on www.ClinicalTrials.gov).

1. Introduction

It is difficult for patients and clinicians to compare the adverse events (AEs) associated with different interventions because AEs are not reported completely in public sources, such as journal articles [1−7]. Nonpublic sources, such as clinical study reports (CSRs), might contain more information [8−11]. Furthermore, multiple sources about clinical trials may contain different information about AEs [8−15].

The methods used to assess and report AEs contribute to challenges for interpretation and synthesis, which previous guidelines have sought to address [16]. Although AEs can be assessed systematically as described in Part 1 of this study [17]—using methods such as those used to assess potential benefits [18,19]—AEs can also be assessed nonsystematically (Table 1) [20,21]. Nonsystematic AEs may be reported by participants to investigators or collected in response to open-ended questions such as ‘‘have you noticed any symptoms since your last visit?’’ AEs may be categorized as ‘‘serious’’ when they lead to hospitalization or death or disrupt normal life functions [22−24].

Terms used to describe AEs are often related (e.g., ‘‘migraine’’ and ‘‘headache’’). Because many AEs are rare, grouping AEs by anatomic or physiological system can increase statistical power and thus increase the possibility of detecting AEs that are caused by interventions. Grouping AEs can also disguise important AEs by combining them with less important AEs (e.g., ‘‘migraine’’ might be more severe than ‘‘headache’’). When information about nonsystematic AEs is submitted to regulators, AEs are often reported in groups (e.g., ‘‘nervous system disorders’’). For submissions to the U.S. Food and Drug Administration (FDA), the Medical Dictionary for Regulatory Activities (MedDRA) [25] classification system replaced COSTART [26] in the late 1990s.

Part 1 and Part 2 of this study about AEs are substudies of the Multiple Data Sources (MUDS) study, which followed a published protocol [27] and protocol amendments [19]. In this substudy, our objectives were to (1) describe the assessment and reporting of nonsystematic AEs identified in eligible clinical trials and (2) compare reporting of AEs in public and nonpublic sources. We assessed nonsystematic AEs (Part 2) apart from systematic AEs (Part 1) 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 controlled 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 (see Table 1) and requested additional information from manufacturers of gabapentin and quetiapine [31]. We included 21 eligible gabapentin trials (80 sources, including six individual patient datasets [IPD]) and seven eligible quetiapine trials (52 sources, including two IPD).

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).

From each source for each trial, we extracted the number of participants who experienced:

(1) One or more nonsystematic AEs;
(2) One or more *serious* nonsystematic AEs (i.e., events that were categorized as ‘‘serious’’ by the trial investigators);
(3) Each nonsystematic AE, which we recorded as the name of the AE (e.g., ‘‘dizziness,’’ ‘‘headache’’);
(4) Each serious nonsystematic AE; and

(5) Each group of AEs (i.e., classified according to anatomic or physiological categories by the trial investigators).

For each AE, we extracted numerical results closest to 8 and 18 weeks. 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). All counts of AEs and results are ''unique,'' meaning that each AE was counted only once, regardless of how many times it appeared across sources (Table 1).

2.3. Reanalysis of IPD

For gabapentin, IPD were available for 6 of 21 (29%) trials [32−37] as Microsoft Access databases. For two quetiapine trials [38,39], we used ABBYY FineReader [40] to extract IPD from tables in an appendix of a CSR that we received as a PDF file; we constructed an electronic database for analysis. Because we did not receive meta-data (e.g., codebooks) for the IPD in either case, we used information in the corresponding CSRs (including incomplete case report forms) to determine the meaning of the variables in the IPD. When both IPD and a CSR were available for a trial, we analyzed the IPD using the methods for handling missing data described in the corresponding CSR, and we compared our results with the CSR to assess agreement.

2.4. Comparison of reporting across multiple data sources

To determine whether *any* AEs were reported in public and nonpublic sources, we compared:

(1) the proportion of public and nonpublic sources that reported nonsystematic AEs; and
(2) the proportion of public and nonpublic sources that reported *serious* nonsystematic AEs.

To determine whether *all* AEs were reported in public and nonpublic sources, we compared:

(1) the proportion of nonsystematic AEs reported in public and nonpublic sources, including the proportion of nonsystematic AEs for which we could calculate the between-arm effect and its precision;
(2) The proportion of *serious* nonsystematic AEs reported in public and nonpublic sources, including the proportion of nonsystematic AEs for which we could calculate the between-arm effect and its precision.
(3) The mean number of AEs reported in public and nonpublic sources; and
(4) The proportion of trials that reported the five most common and 10 randomly selected AEs (Table 2).

To determine whether ''grouped'' AEs were reported in public and nonpublic sources, we compared:

(1) The proportion of public and nonpublic sources that reported AEs for each anatomic or physiological group.

To determine whether public and nonpublic sources included sufficient information about the results, we compared:

(1) the proportion of trials that reported results for each AE that were sufficient to calculate the between-arm effect and its precision (Table 3, Appendix 1).

We performed all analyses using Stata 14 [41].

3. Results

3.1. Public sources reported AEs inconsistently and inadequately

We found one or more public sources for 20 of 21 (95%) gabapentin trials and seven of seven (100%) quetiapine trials. Most public sources did not report the number of participants who experienced one or more AEs (often described as ''any adverse event''); 5 of 20 (25%) gabapentin trials and three of seven (43%) quetiapine trials reported the number of participants who experienced one or more AEs with sufficient information to calculate the between-arm effect and its precision.

By contrast, we found one or more nonpublic sources for 6 of 21 (29%) gabapentin trials and four of seven (57%) quetiapine trials, and all (100%) nonpublic sources reported the number of participants who experienced one or more AEs with sufficient information to calculate the between-arm effect and its precision.

We found both public and nonpublic sources for 5 of 21 (24%) gabapentin trials and four of seven quetiapine (57%) trials. For trials with both public and nonpublic sources, public sources reported few of the AEs we found in nonpublic sources (Fig. 1).

Across all sources for all trials, we found 419 and 471 unique AEs for gabapentin and quetiapine, respectively. Public sources did not report the vast majority of them; 341 of 419 (81%) and 436 of 471 (93%) unique AEs were *not* reported in public sources (Appendix 1). Public sources reported fewer AEs than nonpublic sources for gabapentin (mean [standard deviation] = 3 [6] public, 121 [40] nonpublic) and quetiapine (three [5] public, 159 [158] nonpublic).

Common AEs were relatively likely to appear in public sources, but other AEs were unlikely to appear in public sources (Table 2, Appendix 1). Without access to nonpublic sources, it was unclear whether AEs that were not mentioned in public sources did not occur or whether those AEs occurred and were not reported.

Table 1. Glossary of terms related to adverse events and sources

Term	Definition
Terms related to adverse events (AEs)	
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'' [24]. The U.S. Food and Drug Administration (FDA) and other regulators use this definition [22,23].
Coding Symbols for a Thesaurus of Adverse Reaction Terms (COSTART)	''The Coding Symbols for a Thesaurus of Adverse Reaction Terms (COSTART) is the terminology developed and used by the Food and Drug Administration for the coding, filing and retrieving of post-marketing adverse reaction reports. It provides a method to deal with the variation in vocabulary used by those submit adverse event reports to the FDA'' [26].
Medical Dictionary for Regulatory Activities (MedDRA)	''The Medical Dictionary for Regulatory Activities (MedDRA) Terminology is the international medical terminology developed under the auspices of the International Conference on Harmonisation (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use'' [25]. MedDRA replaced COSTART in the late 1990s [28].
Non-systematic adverse events	According to The Final Rule [20,21], '''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 Source (MUDS) study, a ''result'' is a numerical contrast between a treatment and comparison arm (e.g., relative risk, mean difference).
Serious adverse events	The International Conference on Harmonisation (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'' [24]. FDA and other regulators use this definition [22,23].
Systematic adverse events	According to The Final Rule [20,21], ''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 [29]. For example, ''proportion of participants with 50% change from baseline to 8 weeks on the Young Mania Rating Scale total score.''
Unique adverse event	In the MUDS study, a ''unique'' AE is a defined AE counted only once, regardless of how many times it appeared in all sources for a trial.
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 [30].
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 [30].
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.

3.2. Public sources reported serious AEs inconsistently and inadequately

Among trials with public sources, 4 of 20 (20%) gabapentin trials and four of seven (57%) quetiapine trials reported the number of participants who experienced one or more serious AEs with sufficient information to calculate the between-arm effect and its precision. Among trials with nonpublic sources, six of six (100%) gabapentin trials and four of four (100%) quetiapine trials reported the number of participants who experienced one or more serious AEs with sufficient information to calculate the between-arm effect and its precision.

Across all sources for eligible trials, we found 72 and 46 unique serious AEs in gabapentin and quetiapine

Table 2. Public sources describe common adverse events rather than all adverse events.

Gabapentin for neuropathic pain	Number of trials with results that include sufficient information to calculate a between-arm effect and its precision (*n*)		
	Public sources (*N* = 20 trials) *n* (%)	Nonpublic sources (*N* = 6 trials) *n* (%)	All sources (*N* = 21 trials) *n* (%)
Five most common AEs			
Dizziness	9 (45)	6 (100)	11 (52)
Edema peripheral	5 (25)	6 (100)	9 (43)
Headache	8 (40)	6 (100)	12 (57)
Nausea	8 (40)	6 (100)	12 (57)
Somnolence	9 (45)	6 (100)	11 (52)
Ten randomly selected AEs			
Abdominal pain	1 (5)	6 (100)	6 (29)
Anemia	0 (0)	3 (50)	3 (14)
Chest pain	0 (0)	5 (83)	5 (24)
Constipation	1 (5)	6 (100)	7 (33)
Hernia	0 (0)	2 (33)	2 (10)
Hyperglycemia	0 (0)	6 (100)	6 (29)
Migraine	0 (0)	3 (50)	3 (14)
Skin ulcer	0 (0)	3 (50)	3 (14)
Tooth ache	0 (0)	1 (17)	1 (5)
Upper respiratory tract infection	1 (5)	1 (17)	2 (10)

Quetiapine for bipolar depression	Number of trials with results that include sufficient information to calculate a between-arm effect and its precision		
	Public sources (*N* = 7 trials) *n* (%)	Nonpublic sources (*N* = 4 trials) *n* (%)	All sources (*N* = 7 trials) *n* (%)
Five most common AEs			
Dizziness	6 (86)	4 (100)	6 (86)
Dry mouth	6 (86)	4 (100)	6 (86)
Headache	6 (86)	4 (100)	6 (86)
Sedation	5 (71)	4 (100)	5 (71)
Somnolence	6 (86)	4 (100)	6 (86)
Ten randomly selected AEs			
Dyspepsia	2 (29)	3 (75)	4 (57)
Dystonia	0 (0)	2 (50)	2 (29)
Food craving	0 (0)	2 (50)	2 (29)
Gastritis	0 (0)	1 (25)	1(14)
Hallucination, auditory	0 (0)	2 (50)	2 (29)
Intervertebral disc herniation	0 (0)	1 (25)	1 (14)
Nightmare	0 (0)	2 (50)	2 (29)
Nocturia	0 (0)	1 (25)	1 (14)
Paresthesia	0 (0)	2 (50)	2 (29)
Yawning	0 (0)	1 (25)	1 (14)

Abbreviation: AE, adverse events.

This table includes the five most common AEs (based on data in CSRs) and 10 randomly selected AEs. It shows the number of trials that reported results with sufficient information to calculate the between-arm effect and its precision in public and nonpublic sources. All AEs, including all serious AEs, are listed in Appendix 4 (gabapentin) and Appendix 5 (quetiapine).

Public sources: journal articles, conference abstracts, Food and Drug Administration reviews, trial registrations, and other short reports. Nonpublic sources: clinical study reports (CSRs), CSR-synopses, and individual participant data.

trials, respectively; most of them were reported only in nonpublic sources. In gabapentin and quetiapine trials, 56 of 72 (78%) and 39 of 46 (85%) serious AEs were *not* reported in public sources (Table 3). Public sources described fewer serious AEs than nonpublic sources for gabapentin (mean $\leq$ 1 [1] public, 10 [7] nonpublic) and quetiapine (mean $\leq$ 1 [1] public, 16 [13] nonpublic).

Table 3. Description of adverse event results reported in public and nonpublic sources

	Gabapentin for neuropathic pain			Quetiapine for bipolar depression		
	Public sources	Nonpublic sources	All sources	Public sources	Nonpublic sources	All sources
Number of trials	20 (95%)	6 (29%)	21	7 (100%)	4 (57%)	7
Trials reporting results for the number of participants who experienced one or more AEs with sufficient information to calculate the between-arm effect and its precision	5[a] (25%)	6 (67%)	9[a]	3[a] (43%)	4 (100%)	6[a]
Trials reporting results for the number of participants who experienced one or more *serious* AEs with sufficient information to calculate the between-arm effect and its precision	4[b] (20%)	6 (67%)	9[c]	4[d] (57%)	4 (100%)	5[a]
Number of adverse events						
Unique nonsystematic AEs	68 (16%)	391 (93%)	419	35 (7%)	463 (98%)	471
Nonsystematic AEs with sufficient information to calculate the between-arm effect and its precision	62 (15%)	384 (93%)	411	23 (5%)	459 (99%)	462
Serious nonsystematic AEs	16 (22%)	63 (88%)	72	7 (15%)	46 (100%)	46
Serious nonsystematic AEs with sufficient information to calculate the between-arm effect and its precision	16 (23%)	60 (86%)	70	1 (3%)	38 (100%)	38
AE groups	2 (17%)	12 (100%)	12	4 (15%)	26 (100%)	26

Abbreviation: AE, adverse events.

[a] Additionally, one trial reported results with insufficient information to calculate the between-arm effect and its precision.

[b] Additionally, four trials reported results with insufficient information to calculate the between-arm effect and its precision.

[c] Additionally, three trials reported results with insufficient information to calculate the between-arm effect and its precision.

[d] Additionally, two trials reported results with insufficient information to calculate the between-arm effect and its precision.

Public sources: journal articles, conference abstracts, Food and Drug Administration reviews, trial registrations, and other short reports. Nonpublic sources: clinical study reports (CSRs), CSR-synopses, and individual participant data.

3.3. Public sources did not report AEs grouped by anatomic or physiological system

Where we could identify the system used to code and group AEs, most gabapentin trials used COSTART [26], and most quetiapine trials used MedDRA [25] (Table1).

All CSRs for gabapentin (6 of 6, 100%) and quetiapine trials (2 of 2, 100%) reported AEs grouped by anatomic or physiological system; however, no corresponding public source for those trials reported AEs grouped by anatomic or physiological system. Of the trials reported only in public sources, 2 of 15 (13%) gabapentin and 0 of 3 (0%) quetiapine trials reported AEs by anatomic or physiological system.

When grouped by anatomic or physiological system, some AEs were relatively common among participants in all intervention arms. For example, "Body As A Whole" was common in gabapentin trials and "Psychiatric Disorders" was common in quetiapine trials. There were also clear differences between the drug and placebo arms (Appendices 2 and 3). Compared with placebo, gabapentin was associated with more AEs related to the nervous system (e.g., dizziness, amnesia, somnolence, tremor) and the digestive system (e.g., diarrhea, nausea, vomiting), whereas quetiapine was associated with more AEs related to the nervous system (e.g., convulsions, dizziness, migraine) and gastrointestinal disorders (e.g., abdominal pain, constipation, nausea). Quetiapine was also associated with more AEs related to eye disorders (e.g., diplopia, vision blurred, visual acuity reduced), yet neither specific eye disorders nor grouped eye disorders were described in public sources.

3.4. Public sources were insufficient for synthesis and decision-making

It would be impossible to synthesize information about AEs for all gabapentin or quetiapine trials in a meta-analysis. Using all available sources, a meta-analysis examining the risk of experiencing one or more AEs would include 9 of 21 (43%) gabapentin and 6 of 7 (86%) quetiapine trials (Appendix 4). Using only public sources, the same analysis would include 5 of 21 (24%) gabapentin and 3 of 7 (43%) quetiapine trials (Table 3). We found similar results when examining the number of participants who experienced one or more serious AEs (Table 3).

4. Discussion

In Part 2 of this study, we found that public sources of clinical trials reported little information about which AEs

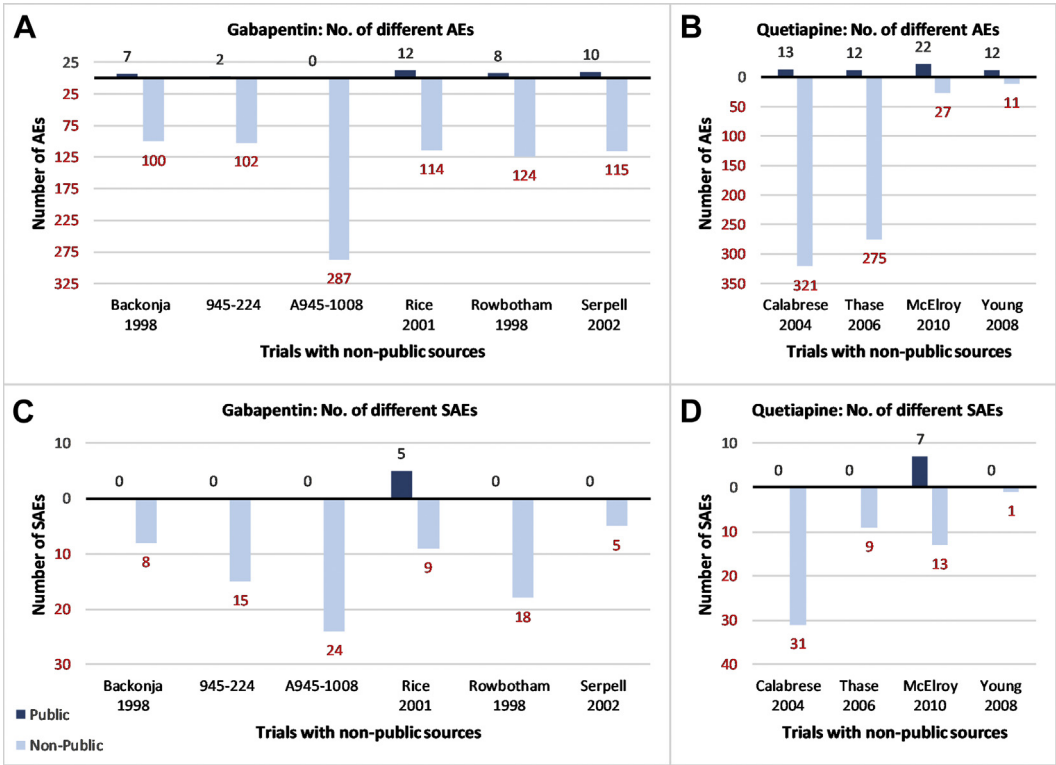

Fig. 1. Fewer nonsystematic adverse events are reported in public sources compared with nonpublic sources: (A) Gabapentin: No. of different AEs in trials with non-public sources. (B) Quetiapine: No. of different AEs in trials with non-public sources. (C) Gabapentin: No. of different SAEs in trials with non-public sources. (D) No. of different SAEs in trials with non-public sources. Public sources: journal articles, conference abstracts, Food and Drug Administration reviews, trial registrations, and other short reports. Nonpublic sources: clinical study reports (CSRs), CSR-synopses, and individual participant data (IPD). This figure includes CSRs and IPD for all gabapentin trials and Calabrese 2004 and Thase 2006 (quetiapine trials). This figure includes only CSR-synopses for McElroy 2010 and Young 2008. AEs, adverse events. SAEs, serious adverse events. No. of AEs, number of AEs reported in each type of source. For example, if a source contained information about "dizziness," "confusion," and "somnolence," the number of adverse events would be 3. No. of SAEs, number of serious AEs reported in each source.

participants experienced and the relative likelihood of experiencing those AEs between arms. It is especially concerning that most serious AEs were not described in public sources. Nonsystematic AEs were grouped for analysis in nonpublic sources, yet public sources almost never reported grouped AEs. For example, FDA prescribing information warns that cataracts have been associated with quetiapine in studies of dogs [42]; we found that eye disorders have been observed in human trials, but no public sources indicated that quetiapine has been associated with eye disorders in people.

All AEs that occur in clinical trials should be reported publicly to prevent reporting bias [43]. That is, systematic reviews and clinical guidelines that rely on public information about nonsystematic AEs might miss important AEs and include systematically biased information about other AEs. If public reports of clinical trials include only common AEs, and if public reports include only evidence that confirms certain AEs are related to interventions, then both rare AEs and disconfirming evidence will be overlooked by decision-makers. Although decision-makers might be overwhelmed [44] if information about hundreds of nonsystematic AEs were reported in journal articles and other summaries of clinical trials (e.g., Appendices 4 and 5) [45,46], information about all nonsystematic AEs could be reported in public registries such as www.ClinicalTrials.gov.

Most trials are not designed to compare individual nonsystematic AEs between arms. Grouping AEs by anatomic or physiological category can increase statistical power, and it can limit the number of comparisons and the potential for false positives in clinical trials [47]. Grouping might also allow investigators to hide certain AEs by combining important AEs with less important AEs. For example, "headache" and "tremor" are both "nervous system disorders" that can be important to patients; the latter might be more likely to be important and long-lasting. To prevent reporting bias, methods for grouping and analyzing nonsystematic AEs should be described in trial protocols and statistical analysis plans.

Because nonsystematic AEs are not collected using specific tools or instruments, nonsystematic AEs might be assessed inconsistently across trials or even within a trial by investigators who use different procedures to assess an outcome. Although it might be impossible to anticipate which AEs will occur in the first clinical trial of a new intervention, AEs could be identified for systematic collection based on

preclinical studies, trials of related interventions (e.g., drugs in the same class), observational studies, and studies of the same intervention for other indications [46]. As described in Part 1, systematic collection of anticipated and patient-important AEs would promote greater consistency across trials and across sites within individual trials [17].

This study was limited to two drugs and health conditions, and eligible trials were relatively short in duration. Additional research is needed to assess whether our findings apply to other interventions and longer trials in which more AEs and more serious AEs occur. Our findings were similar for both cases, and our findings are consistent with previous studies that compared public and nonpublic sources [8−11]; although the clinical implications of underreporting nonsystematic AEs vary across interventions and conditions, our methodological findings might be generalizable beyond these two drugs and conditions.

''Open science'' and data sharing offer partial solutions to the problems identified in this study. In future studies, investigators could report all AEs in trial registries and through data sharing services such as Vivli (https://vivli.org/). Many commonly prescribed medications were evaluated and approved by regulators before trial registration and reporting requirements were commonplace [20,48]; nonetheless, it is important for stakeholders to know the AEs associated with interventions in current use, regardless of when those interventions were evaluated in clinical trials. Standards are needed for collecting, summarizing, sharing, and synthesizing AEs in clinical trials [17], including standards for sharing the results of ''legacy'' studies [49].

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

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.2019.04.020.

References

[1] Breau RH, Gaboury I, Scales CD, Fesperman SF, Watterson JD, Dahm P. Reporting of harm in randomized controlled trials published in the urological literature. J Urol 2010;183(5):1693−7.

[2] Carlesso LC, MacDermid JC, Santaguida LP. Standardization of adverse event terminology and reporting in orthopaedic physical therapy: application to the cervical spine. J Orthop Sport Phys 2010;40(8):455−63.

[3] de Vries TW, van Roon EN. Low quality of reporting adverse drug reactions in paediatric randomised controlled trials. Arch Dis Child 2010;95:1023−6.

[4] Hochman M, McCormick D. Characteristics of published comparative effectiveness studies of medications. JAMA 2010;303:951−8.

[5] Loke YK, Derry S. Reporting of adverse drug reactions in randomised controlled trials−a systematic survey. BMC Clin Pharmacol 2001;1(1):1−6.

[6] Pitrou I, Boutron I, Ahmad N, Ravaud P. Reporting of safety results in published reports of randomized controlled trials. Arch Intern Med 2009;169:1756−61.

[7] Pope A, Adams C, Paton C, Weaver T, Barnes TR. Assessment of adverse effects in clinical studies of antipsychotic medication: survey of methods used. Br J Psychiatry 2010;197(1):67−72.

[8] 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.

[9] Jefferson T, Jones MA, Doshi P, Del Mar CB, Heneghan CJ, Hama R, et al. Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children. Cochrane Database Syst Rev 2012;1:CD008965.

[10] Le Noury J, Nardo JM, Healy D, Jureidini J, Raven M, Tufanaru C, et al. Restoring study 329: efficacy and harms of paroxetine and imipramine in treatment of major depression in adolescence. BMJ 2015;351:h4320.

[11] Rodgers MA, Brown JV, Heirs MK, Higgins JP, Mannion RJ, Simmonds MC, 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.

[12] Hart B, Lundh A, Bero L. Effect of reporting bias on meta-analyses of drug trials: reanalysis of meta-analyses. BMJ 2012;344:d7202.

[13] Halfpenny NJA, Quigley JM, Thompson JC, Scott DA. Value and usability of unpublished data sources for systematic reviews and network meta-analyses. Evid Based Med 2016;21(6):208−13.
[14] Hartung DM, Zarin DA, Guise J-M, McDonagh M, Paynter R, Helfand M. Reporting discrepancies between the ClinicalTrials. gov results database and peer-reviewed publications. Ann Intern Med 2014;160:477−83.
[15] Riveros C, Dechartres A, Perrodeau E, Haneef R, Boutron I, Ravaud P. Timing and completeness of trial results posted at ClinicalTrials. gov and published in journals. PLoS Med 2013;10(12): e1001566.
[16] 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.
[17] Mayo-Wilson E, Fusco N, Li T, Hong H, Canner JK, Dickersin K, et al. Harms are assessed inconsistently and reported inadequately Part 1: systematic adverse events. J Clin Epidemiol 2019;113:20−7.
[18] 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.
[19] 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.
[20] 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.
[21] Zarin DA, Tse T, Williams RJ, Carr S. Trial reporting in ClinicalTrials. gov—the final rule. N Engl J Med 2016;375:1998−2004.
[22] 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.
[23] 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.
[24] 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.
[25] International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use. In: Introductory guide MedDRA version 17.0. Available at https://www.meddra.org/sites/default/files/guidance/file/intguide_17_0_english.pdf. Accessed May 12, 2018.
[26] U.S. Food and Drug Administration (FDA). Coding symbols for a thesaurus of adverse reaction terms. 5th ed. 1995. Available at https://bioportal.bioontology.org/ontologies/COSTART. Accessed May 12, 2018.
[27] 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.
[28] MedDRA - Medical dictionary for regulatory activities. 1999. Available at https://www.meddra.org/about-meddra/history. Accessed May 12, 2018.
[29] 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.
[30] 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.
[31] Mayo-Wilson E, Doshi P, Dickersin K. Are manufacturers sharing data as promised? BMJ 2015;351:h4169.
[32] Parke-Davis. Research report: 720-03908. A double-blind placebo-controlled trial of gabapentin for treatment of painful diabetic peripheral neuropathy 1998: Protocol Number: 945-210.
[33] Parke-Davis. Research report: 995-00070. Double-blind, randomized, placebo-controlled, parallel groups, multi-center trial to determine the efficacy and safety of Neurontin (gabapentin) in subjects with peripheral neuropathy (post-herpetic neuralgia) 1998: Protocol Number: 945-211.
[34] Parke-Davis. Research report: 720-04130. A double-blind placebo-controlled trial with 3 doses of gabapentin for treatment of painful diabetic neuropathy 2000: Protocol Number: 945-224.
[35] Parke-Davis. Research report: 430-00125. A double blind placebo controlled trial of gabapentin for the treatment of patients exhibiting symptoms of neuropathic pain 2000: Protocol Number: 945-430-306 (NN 026).
[36] Parke-Davis. Research report: 430-00124. A double-blind placebo controlled trial of gabapentin for the treatment of post herpetic neuralgia 2000: Protocol Number: 945-430-295 (NN 025).
[37] Pfizer. Final study report: A 15 week randomized, double-blind, placebo controlled, parallel-group, multicenter study of Neurontin (gabapentin) for efficacy and quality of life in patients with painful diabetic peripheral neuropahty 2005: Protocol A945-1008.
[38] 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 2005: Study code: 5077US/0049.
[39] AstraZeneca. A confirmatory multicenter, double-blind, randomized, placebo-controlled study of the use of quetiapine fumarate (SEROQUEL) in the treatment of patients with bipolar depression [D1447C00135] 2005.
[40] ABBYY. FineReader.. 11th ed. Milpitas, CA: ABBYY; 2011.
[41] StataCorp LP. Stata Statistical Software: Release 14 2015: College Station, TX.
[42] Food and Drug Administration. Highlights of prescribing information: Seroquel (R) (quetiapine fumarate) tablets, for oral use (Reference ID: 4060088). 2017. Available at https://www.accessdata.fda.gov/drugsatfda_docs/label/2017/020639s065lbl.pdf. Accessed May 12, 2018.
[43] Dickersin K, Chalmers I. Recognizing, investigating and dealing with incomplete and biased reporting of clinical research: from Francis Bacon to the WHO. J R Soc Med 2011;104(12):532−8.
[44] Sivanathan N, Kakkar H. The unintended consequences of argument dilution in direct-to-consumer drug advertisements. Nat Hum Behav 2017;1(11):797.
[45] Mayo-Wilson E, Fusco N, Li T, Hong H, Canner JK, Dickersin K. Opportunities for selective reporting of non-systematic adverse events in randomized controlled trials: an observational study 2018: Submitted.
[46] 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.
[47] 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.
[48] DeAngelis CD, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, et al. Clinical trial registration: a statement from the International Committee of Medical Journal. JAMA 2004;292:1363−4.
[49] 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.