ELSEVIER
Journal of Clinical Epidemiology xxx (xxxx) xxx
Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Reporting transparency and completeness in trials: Paper 3 – trials conducted using administrative databases do not adequately report elements related to use of databases

Mahrukh Imran[a], Kimberly Mc Cord[b], Stephen J. McCall[c,d], Linda Kwakkenbos[e], Margaret Sampson[f], Ole Fröbert[g], Chris Gale[h], Lars G. Hemkens[b], Sinéad M Langan[i], David Moher[j], Clare Relton[k], Merrick Zwarenstein[l,m], Edmund Juszczak[c,n], Brett D. Thombs[a,o,p,q,r,s,t,*], on behalf of CONSORT Extension for Trials Conducted Using Cohorts and Routinely Collected Data Group[1,t]

[a] *Lady Davis Institute for Medical Research, Jewish General Hospital, 4333 Cote Ste. Catherine Road, Montréal, Quebec, Canada*
[b] *Basel Institute for Clinical Epidemiology and Biostatistics, Department of Clinical Research, University Hospital Basel, University of Basel, Basel, Switzerland*
[c] *National Perinatal Epidemiology Unit Clinical Trials Unit, Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom*
[d] *Center for Research on Population and Health, Faculty of Health Sciences, American University of Beirut, Ras Beirut, Lebanon*
[e] *Behavioural Science Institute, Clinical Psychology, Radboud University, Nijmegen, the Netherlands*
[f] *Library Services, Children's Hospital of Eastern Ontario, Ottawa, Ontario, Canada*
[g] *Department of Cardiology, Faculty of Health, Örebro University, Örebro, Sweden*
[h] *Neonatal Medicine, School of Public Health, Faculty of Medicine, Imperial College London, London, United Kingdom*
[i] *Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom*
[j] *Centre for Journalology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada*
[k] *Centre for Clinical Trials and Methodology, Barts Institute of Population Health Science, Queen Mary University, London, United Kingdom*
[l] *Department of Family Medicine, Western University, London, Ontario, Canada*
[m] *Institute for Clinical Evaluative Sciences, Toronto, Ontario, Canada*
[n] *Nottingham Clinical Trials Unit, University of Nottingham, University Park, Nottingham, United Kingdom*
[o] *Department of Psychiatry, McGill University, Montreal, Quebec, Canada*
[p] *Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Quebec, Canada*
[q] *Department of Medicine, McGill University, Montreal, Quebec, Canada*
[r] *Department of Psychology, McGill University, Montreal, Quebec, Canada*
[s] *Department of Educational and Counselling Psychology, McGill University, Montreal, Quebec, Canada*
[t] *Biomedical Ethics Unit, McGill University, Montreal, Quebec, Canada*

Accepted 7 September 2021; Available online xxx

Abstract

Objective: We evaluated reporting completeness and transparency in randomized controlled trials (RCTs) conducted using administrative data based on 2021 CONSORT Extension for Trials Conducted Using Cohorts and Routinely Collected Data (CONSORT-ROUTINE) criteria.

Study Design and Setting: MEDLINE and the Cochrane Methodology Register were searched (2011 and 2018). Eligible RCTs used administrative databases for identifying eligible participants or collecting outcomes. We evaluated reporting based on CONSORT-ROUTINE, which modified eight items from CONSORT 2010 and added five new items.

Results: Of 33 included trials (76% used administrative databases for outcomes, 3% for identifying participants, 21% both), most were conducted in the United States (55%), Canada (18%), or the United Kingdom (12%). Of eight items modified in the extension; six were adequately reported in a majority (>50%) of trials. For the CONSORT-ROUTINE modification portion of those items, three items were reported adequately in >50% of trials, two in <50%, two only applied to some trials, and one only had wording modifications and was not evaluated. For five new items, four that address use of routine data in trials were reported inadequately in most trials.

All authors declare no competing interests.

[1] CONSORT Extension for Trials Conducted Using Cohorts and Routinely Collected Data Group: Eric I. Benchimol; Isabelle Boutron; Marion K. Campbell; David Erlinge; John Fletcher; Jon Nicholl; Philippe Ravaud; Danielle B. Rice; Maureen Sauvé; Lehana Thabane; David Torgerson; Rudolf Uher; Helena M. Verkooijen.

* Corresponding author. Tel.: 514 340-8222/ex25112.
E-mail address: brett.thombs@mcgill.ca (B.D. Thombs).

https://doi.org/10.1016/j.jclinepi.2021.09.010

Conclusion: How administrative data are used in trials is often sub-optimally reported. CONSORT-ROUTINE uptake may improve reporting.

Key Words: Administrative data; CONSORT; CONSORT-ROUTINE; Randomized controlled trials; Reporting guideline; Routinely collected data

What is new?

Key Findings

- Among items modified from the 2010 CONSORT statement, items on describing the use of an administrative database in the abstract (91%), including the administrative dataset in the statement of trial design (82%), and describing the source of outcome data (88%) were adequately reported in most trials; modifications related to how the use of administrative data may have influenced generalizability (21%) and funding of the database (6%) were not reported adequately in most trials.
- New CONSORT-ROUTINE items on eligibility criteria for inclusion in the administrative database (6% adequate, 21% partially adequate), description of record linkages (3%, 33%), listing of codes and adjudication of outcomes (0%, 15%), and providing a full description of the administrative database (9%, 82%) were not reported adequately in most trials.

What this study adds to what was known?

- No previous studies have examined completeness and transparency of reporting of recent randomized controlled trials conducted using administrative databases published prior to the development of the CONSORT-ROUTINE statement.

What is the implication and what should change now?

- The way in which administrative data are used in trials is often not reported adequately and may reduce utility of published trial reports.
- Authors should refer to the 2021 CONSORT Extension for Trials Conducted Using Cohorts and Routinely Collected Data (CONSORT-ROUTINE) for guidance on reporting of trials conducted using cohorts, registries, electronic health records, and administrative databases.

1. Introduction

There is growing interest in the use of administrative databases to evaluate health care interventions [1]. Health system administrative databases include information collected for administrative or billing purposes (e.g., Medicare data in the United States) that is routinely collected during clinic, hospital, laboratory, or pharmacy visits. These data can provide a readily available source of "real-world" data on a large population over expansive geographic regions [2]. Administrative databases are increasingly accessible to researchers and are being more frequently utilized in randomized controlled trials (RCTs) as an inexpensive and reliable resource of data at multiple stages of trials, from identifying and recruiting eligible participants to determining study outcomes [3,4].

There are several possible advantages of using administrative data to conduct RCTs, such as more efficient identification and recruitment of participants, improved data collection and outcome ascertainment, and improved feasibility due to reductions in cost, time, and resources [5]. However, several factors must be considered in these types of RCTs. For instance, the accuracy of administrative data and potential for bias should be taken into account if complete data are not available for all potential trial participants. Many large administrative databases have been developed by governments and private insurers, primarily for financial and administrative purposes, rather than clinical research, and therefore vary in completeness and accuracy [3,6,7]. Characteristics of participants in an administrative database used to select trial participants and how well they match the true target population for the trial should be taken into consideration because the representativeness of trial participants is dependent on that of the administrative database. In addition, there may be unique challenges in linking administrative data to other sources of data, stemming, for example, from linkage errors when records cannot be linked or are linked incorrectly [8].

The CONsolidated Standards of Reporting Trials (CONSORT) 2010 reporting guideline, which includes a 25-item checklist and flow diagram, was developed to improve the quality of reporting of parallel group RCTs [9]. Several extensions of the CONSORT Statement have been developed to encourage better reporting of alternative trial designs, including multiarm parallel group randomized trials [10], cluster trials [11], pilot and feasibility trials [12], and pragmatic trials [13], for example. CONSORT-ROUTINE, which was published in 2021, was developed as an extension for trials conducted using cohorts and routinely collected data, including registries, electronic health records, and administrative data, and provides a minimal set of items that should be included in reports of these types

Please cite this article as: M. Imran et al., Reporting transparency and completeness in trials: Paper 3 – trials conducted using administrative databases do not adequately report elements related to use of databases, Journal of Clinical Epidemiology, https://doi.org/10.1016/j.jclinepi.2021.09.010

of trials [14]. CONSORT-ROUTINE was needed because, although RCTs conducted using cohorts and routinely collected data share elements with two-arm parallel groups RCTs covered in the CONSORT 2010 statement, there are aspects that differ and require additional or modified reporting elements.

The present review examines RCTs identified as part of a broader scoping review [15] that was conducted to support the development of CONSORT-ROUTINE [14]. We aimed to (1) describe characteristics of RCTs conducted using administrative data and published after the CONSORT 2010 statement; and (2) assess and describe the quality of reporting of trials using administrative data by coding the completeness and transparency of all newly added and modified items from CONSORT-ROUTINE. For modified items, we also evaluated the transparency and completeness of reporting of the CONSORT 2010 items to determine if any suboptimal reporting was specific to the extension or if reporting was deficient even based on the CONSORT 2010 checklist item available at the time of publication. Since CONSORT-ROUTINE was published in 2021, the present study serves as a benchmark for pre-CONSORT-ROUTINE reporting of trials conducted using administrative databases.

2. Methods

The study protocol is accessible via the Open Science Framework: https://osf.io/dp23x/.

2.1. Inclusion and exclusion criteria for RCTs using administrative databases

The main scoping review included reports of trials that had used cohorts or routinely collected data to both identify or screen for participants and ascertain trial outcomes, as well as protocols, commentaries, and reviews of methodological aspects of conducting trials using cohorts or routinely collected data [15]. For the present review, eligible RCTs had to have used an administrative database to: (1) identify potentially eligible participants for the trial; (2) ascertain trial outcomes; or (3) both. Administrative databases were defined as databases not originally intended for research that are used for routine governance and program administration. Some examples include public or private insurance databases, birth or death registries, or employment and social care databases.

Methodological reviews, commentaries, and trial protocols were excluded. Publications that reported cost-effectiveness studies or RCTs assessing non-health outcomes were also excluded. Although the main scoping review searched for publications from 2007 to 2018, we restricted the present review to trials published from 2011 to 2018 to include only those published following the publication of the CONSORT 2010 statement.

2.2. Search strategy and study selection

Ovid MEDLINE Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE Daily and Ovid MEDLINE and EBM Reviews – Cochrane Methodology Registry (Final issue, third Quarter 2012) were searched from January 2007 to March 2018 (Cochrane Methodology Register up to last update in July 2012). Search strategies were developed by an experienced research librarian familiar with knowledge synthesis related to research methods and reporting with input from the project team and were peer reviewed using the Peer Review of the Electronic Search Strategy (PRESS) [16]. Appendix 1 provides search terms used to identify RCTs conducted using administrative data. References were imported into Refworks, and duplicates were removed. References were then imported into the systematic review software DistillerSR (Evidence Partners, Ottawa, Canada) [17]. The coding manual for inclusion and exclusion is shown in Appendix 2.

Titles and abstracts were screened independently by two reviewers. A liberal accelerated method, where titles and abstracts are screened by one reviewer and excluded publications are screened by a second reviewer, was used to identify publications for inclusion for full text review [18]. This was done in random order so that reviewers were blind to whether the other reviewer had already made a decision on any given title and abstract. Any trial that appeared potentially eligible was selected for full-text review, even if administrative database use was not described explicitly in the abstract. Full texts were screened independently by two reviewers, and any disagreements were resolved by discussion and consensus with involvement of a third reviewer, if necessary.

2.3. Data extraction

Data were extracted from all identified studies into a predefined form. Items extracted from each RCT publication included: research question of the trial, level of randomisation (cluster, individual), setting, disease of interest, use of administrative database (participant identification, trial data collection), intervention (surgical, screening, drug, other), comparator (placebo, active comparison, usual care), primary outcome, whether primary outcome was assessed using the administrative database, country where the RCT was conducted, and the number of clusters or participants randomized. These items were presented for all trials and separately by cluster RCTs and individually randomized RCTs. We also classified studies into reports of primary or secondary trial outcomes to evaluate any differences in the quality of reporting between primary and secondary reports. *Primary publications* were defined as reports on the trial's primary outcome(s) and also, possibly, other trial outcomes. *Secondary publications* were defined as reports on only secondary outcomes or other

Please cite this article as: M. Imran et al., Reporting transparency and completeness in trials: Paper 3 – trials conducted using administrative databases do not adequately report elements related to use of databases, Journal of Clinical Epidemiology, https://doi.org/10.1016/j.jclinepi.2021.09.010

post-hoc outcomes; reports that described reporting secondary outcomes or that referred to a previous publication of trial outcomes were coded as secondary reports.

Data were extracted by one investigator and validated by a second investigator.

2.4. Evaluation of completeness and transparency of reporting

We evaluated the completeness and transparency of all items in CONSORT-ROUTINE that were either new items (N = 5) or were items from the CONSORT 2010 statement [14] that were modified (N = 8). For modified items, we evaluated reporting both based on the original CONSORT 2010 items and based on the modified portion of the items. We did this in order to determine if any suboptimal reporting was related to inadequate reporting based on the original CONSORT 2010 checklist item, which was available at the time of publication of the included trials, or to the item modification. We did not evaluate reporting of items that were unmodified from the CONSORT 2010 statement.

For each included trial, reporting of each item was categorized as 'adequately reported', 'partially reported', 'inadequately or not reported', or 'not applicable'. A coding manual was devised to ensure consistent assessment of reporting (see Appendix 3). This manual was also used in separate studies that assessed the completeness and transparency of reporting in registries and electronic health records [19,20]. The data extraction rules and coding manual were pilot tested in five RCTs by four investigators to clarify wording and calibrate agreement between reviewers. The assessment of completeness and transparency of reporting was then conducted by one reviewer and validated by a second reviewer. Any disagreements were resolved by discussion and consensus with a third reviewer consulted as necessary. Results were synthesized by totalling the number and percentage of studies adequately, partially, and inadequately or not applicable for each item.

3. Results

We retrieved 660 unique citations from the electronic database search, of which 509 were excluded after title and abstract review and 118 after full-text review, leaving 33 publications for data extraction and quality assessment. See Figure 1. References for all includes studies are in Appendix 4.

3.1. Characteristics of included RCTs

Of the 33 included studies, 25 (76%) were primary publications, and eight (24%) were secondary publications; 20 (61%) were individually randomized, and 13 (39%) were cluster RCTs. There were 25 (76%) that used administrative databases to assess outcomes only, seven (21%) that used them for both participant identification and outcome assessment, and one (3%) that used them for identification of participants only.

Most trials were performed in the United States (N = 18, 55%), followed by Canada (N = 6, 18%) and the United Kingdom (N = 4, 12%). The interventions most frequently tested were educational (N = 10, 30%), multicomponent (N = 7, 21%), and drugs (N = 4, 12%). Comparators included usual care (N = 25, 76%) and alternative therapies (N = 8, 24%). Commonly reported primary outcomes were mortality (N = 5, 15%), hospitalization (N = 5, 15%), and surrogate outcomes (N = 4, 12%). Of the 33 included studies, 22 (67%) used the administrative database for ascertaining the primary trial outcome and 10 (30%) for ascertaining secondary outcomes; for one trial (3%) it was unclear whether primary or secondary outcomes were ascertained (see Table 1 and Appendix 5 for table by cluster versus individually randomized trials).

3.2. Baseline assessment of completeness and transparency of reporting

Results for all included trials are available at https://osf.io/hs9tz/.

3.2.1. CONSORT 2010 items with modifications in CONSORT-ROUTINE

Eight CONSORT 2010 items were modified in CONSORT-ROUTINE. As shown in Table 2, the original version of six of these items ("Structured summary" (88%), "Eligibility criteria" (85%), "Outcome definition" (94%), "Participant flow" (67%), "Interpretation" (97%) and "Funding" (58%)) were adequately reported in a majority of trials (Table 2). Item "Trial design" was adequately reported in 39%, and Item "Allocation concealment mechanism" was adequately reported in 27%. Compliance to the CONSORT 2010 criteria was generally similar in primary and secondary publications (see Appendix 6).

In the modified portions of the modified items, three items were adequately reported in a majority of trial publications; ("Modified – Administrative database use and name in the abstract" (91%), "Modified – Description of trial design" (82%) and "Modified – Outcomes" (88%)). One item "Modified – Funding" was adequately reported for only 6% but partially reported for 61%. Another, "Modified – Interpretation of results", was reported adequately in only 21%. The remaining two items were not applicable for assessment in a majority of trials because the trials used administrative data for assessing outcomes only, but not for identifying eligible participants or as a mechanism for allocating participants to trial arms: ("Modified – Eligibility criteria for participants" (82%) and "Modified – Participant flow" (84%)). Item "Modified – Allocation concealment" was not coded separately as the modification was a clarification of the original item. Results were

Please cite this article as: M. Imran et al., Reporting transparency and completeness in trials: Paper 3 – trials conducted using administrative databases do not adequately report elements related to use of databases, Journal of Clinical Epidemiology, https://doi.org/10.1016/j.jclinepi.2021.09.010

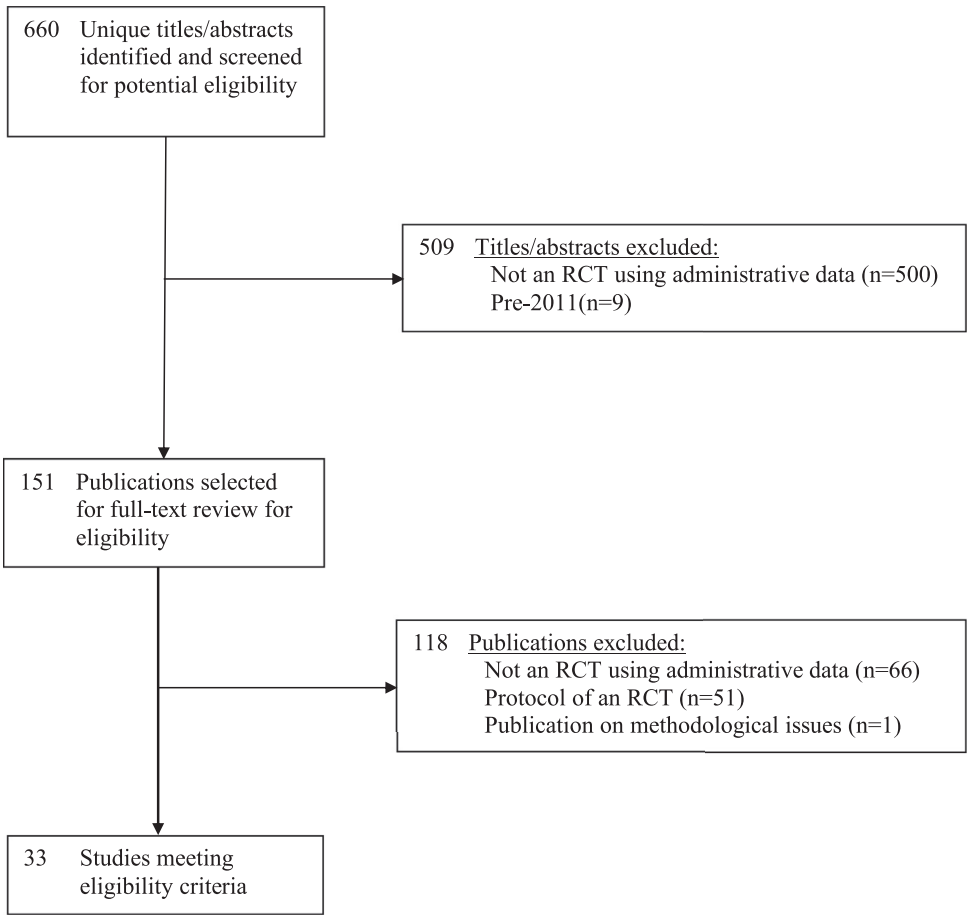

Figure 1. Flow diagram of publication selection process – randomized controlled trials conducted using administrative data

similar when stratified by primary and secondary publication type (Appendix 6).

3.2.2. New items in CONSORT-ROUTINE

Of the five new items evaluated, four items were inadequately reported in >50% of trials; "Eligibility (for cohort or routinely collected database)" (73%), "Description of record linkage" (64%) and "List of codes, monitoring and adjudication for outcomes" (82%). Item "Description of the cohort or routinely collected database" was adequately reported in only 9% but partially reported in 82%. Only one item "Informed consent" (79%) was adequately reported in most of the trials.

4. Discussion

We evaluated the degree to which 33 RCTs conducted using administrative data reported results consistent with existing CONSORT reporting criteria and with new criteria in CONSORT-ROUTINE [14]. Among eight modified items, seven included additional content in the modification. Based on the CONSORT 2010 versions of the eight items, six items related to elements of trial design, interpretation, and funding were adequately reported in at least 50% of included trials, but two items related to randomisation and allocation methodology were not typically reported adequately. Considering only the modified parts of the seven items with additional content, three items related to describing that routinely collected data were used in the abstract, including the administrative dataset in the statement of the trial design, and describing the source of outcome data were adequately reported in a majority of the trials. Modifications related to interpreting how the use of routinely collected data may have influenced the trial or its generalizability and reporting funding of the routinely collected database were not reported adequately in most trials. Two items with modifications were not evaluated in most trials because they were only applicable to trials that used administrative databases for purposes other than assessing outcomes (e.g., eligibility, recruitment, allocation). Among the five new items, four related to aspects of using the routinely collected data were not reported adequately in most trials, whereas one item that requires reporting of aspects of consent was adequately reported in more than 50% of trials.

Among key reporting gaps, most studies did not adequately describe the administrative database used in the RCT, which is important for assessing the validity of the

Please cite this article as: M. Imran et al., Reporting transparency and completeness in trials: Paper 3 – trials conducted using administrative databases do not adequately report elements related to use of databases, Journal of Clinical Epidemiology, https://doi.org/10.1016/j.jclinepi.2021.09.010

Table 1. Characteristics of trials conducted using administrative databases

	Total (%) (n = 33)
Primary publication (versus secondary)	25 (76%)
Use of administrative data in trial	
Identification of patients	1 (3%)
Outcome ascertainment	25 (76%)
Both identification and outcomes	7 (21%)
Administrative data used for primary outcome (versus no or unclear)	22 (67%)
Setting	
Inpatient	11 (33%)
Primary care	10 (30%)
Other[I]	12 (36%)
Country	
USA	18 (55%)
Canada	6 (18%)
UK	4 (12%)
Other[II]	2 (6%)
Disease type	
General health	12 (36%)
Cardiovascular disease	9 (27%)
Other[III]	12 (36%)
Intervention	
Educational	10 (30%)
Multicomponent	7 (21%)
Drug	4 (12%)
Other[IV]	12 (36%)
Active comparator (versus usual care)	8 (24%)
Primary outcome	
Mortality	5 (15%)
Hospitalization	5 (15%)
Surrogate	4 (12%)
Other[V]	19 (58%)
Sample size	
Clusters (Median and IQR) in 13 cluster randomised trials	101 [73–221]
Participants (Median and IQR) in 13 cluster randomised trials	119,910 [86,998–526,850]
Participants (Median and IQR) in 20 individually randomised trials	32,804 [32,804–33,081]

[I] Community medicine, outpatient, residential setting, multiple settings.
[II] Europe, Australia, India, New Zealand.
[III] Mental health, respiratory disease, diabetes, cancer, potentially inapproproate medicines, drug side effects, infection, disability, homelessness.
[IV] Guideline/reminder-based, elephone/web-based care, Family Finding program, referral, housing, health care provider support, surgical.
[V] Self-reported, insurance claims, uptake of treatment, disease occurence, no primary outcome, adherence, risk of injury, multiple/composite outcomes, injury rate.

data used and may have implications for trial generalizability. Information related to database eligibility criteria was also inadequately reported, which could negatively affect the ability of readers to judge the representativeness of the database to the population targeted for the RCT intervention. Details on linkage methodology between databases, which can add biases due to incomplete or incorrect matching of participants, was also poorly reported

Please cite this article as: M. Imran et al., Reporting transparency and completeness in trials: Paper 3 – trials conducted using administrative databases do not adequately report elements related to use of databases, Journal of Clinical Epidemiology, https://doi.org/10.1016/j.jclinepi.2021.09.010

Table 2. Completeness and transparency of reporting for CONSORT 2010 items that were modified, modified items, and new items in CONSORT-ROUTINE[I]

	Item[II]	CONSORT 2010 Item that was modified	CONSORT-ROUTINE item text	N = 33			
				Adequately reported N (%)	Partially reported N (%)	Inadequately or not reported N (%)	Not applicable N (%)
Title and abstract							
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts).		29 (88%)	4 (12%)	0 (0%)	-
			Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts). **Specify that a cohort or routinely collected data were used to conduct the trial and, if applicable, provide the name of the cohort or routinely collected database(s) (Modified)**	30 (91%)	3 (9%)	0 (0%)	-
Methods							
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio		11 (33%)	9 (27%)	13 (39%)	-
			Description of trial design (such as parallel, factorial) including allocation ratio, **that a cohort or routinely collected database(s) was used to conduct the trial (such as electronic health record, registry) and how the data were used within the trial (such as identification of eligible trial participants, trial outcomes) (Modified)**	27 (82%)	6 (18%)	0 (0%)	-
Cohort or routinely collected database	ROUTINE-1		Name, if applicable, and description of the cohort or routinely collected database(s) used to conduct the trial, including information on the setting (such as primary care), locations, and dates, (such as periods of recruitment, follow-up, and data collection) (New)	3 (9%)	27 (82%)	3 (9%)	-
	ROUTINE-2		Eligibility criteria for participants in the cohort or routinely collected database(s) (New)	2 (6%)	7 (21%)	24 (73%)	-
	ROUTINE-3		State whether the study included person-level, institutional-level, or other data linkage across two or more databases and, if so, linkage techniques and methods used to evaluate completeness and accuracy of linkage (New)	1 (3%)	11 (33%)	21 (64%)	-

(*continued on next page*)

Table 2 (*continued*)

	Item[ll]	CONSORT 2010 Item that was modified	CONSORT-ROUTINE item text	N = 33			
				Adequately reported N (%)	Partially reported N (%)	Inadequately or not reported N (%)	Not applicable N (%)
Trials participants	4a	Eligibility criteria for participants		28 (85%)	4 (12%)	1 (3%)	-
			Eligibility criteria for trial participants, **including information on how to access the list of codes and algorithms used to identify eligible participants, information on accuracy and completeness of data used to ascertain eligibility, and methods used to validate accuracy and completeness (e.g., monitoring, adjudication), if applicable (Modified)**	0 (0%)	5 (15%)	1 (3%)	27 (82%)
	ROUTINE-4		Describe whether and how consent was obtained (New)	26 (79%)	1 (3%)	6 (18%)	-
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed		31 (94%)	2 (6%)	0 (0%)	-
			Completely defined prespecified primary and secondary outcome measures, including how and when they were ascertained and **the cohort or routinely collected database(s) used to ascertain each outcome (Modified)**	29 (88%)	4 (12%)	0 (0%)	0 (0%)
	ROUTINE-5		Information on how to access the list of codes and algorithms used to define or derive the outcomes from the cohort or routinely collected database(s) used to conduct the trial, information on accuracy and completeness of outcome variables, and methods used to validate accuracy and completeness (e.g., monitoring, adjudication), if applicable (New)	0 (0%)	5 (15%)	27 (82%)	1 (3%)
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Mechanism used to implement the random allocation sequence **(such as embedding an automated randomizer within the cohort or routinely collected database(s)),** describing any steps taken to conceal the sequence until interventions were assigned (Modified)	9 (27%)	3 (9%)	21 (64%)	-

(*continued on next page*)

Table 2 (*continued*)

	Item[II]	CONSORT 2010 Item that was modified	CONSORT-ROUTINE item text	N = 33			
				Adequately reported N (%)	Partially reported N (%)	Inadequately or not reported N (%)	Not applicable N (%)
Results							
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome		22 (67%)	9 (27%)	2 (6%)	-
			For each group, the number of participants **in the cohort or routinely collected database(s) used to conduct the trial and the numbers screened for eligibility,** randomly assigned, **offered and accepted interventions (e.g., cohort multiple RCTs),** received intended treatment, and analysed for the primary outcome (Modified)	1 (3%)	5 (15%)	1 (3%)	26 (84%)
Discussion							
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		32 (97%)	1 (3%)	0 (0%)	-
			Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence, **including the implications of using data that were not collected to answer the trial research questions (Modified)**	7 (21%)	1 (3%)	25 (76%)	-
Other information							
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders		19 (58%)	13 (39%)	1 (3%)	-
			Sources of funding and other support **for** both the trial and **the cohort or routinely collected database(s), role of funders (Modified)**	2 (6%)	20 (61%)	11 (33%)	-

[I] For modified items, modifications are shown in bold. For those items, only portion modified was evaluated.
[II] Item numbers reflect numbers in original 2010 CONSORT checklist that were modified or new items. New items are designated by "CONSORT-ROUTINE".

in a majority of the trials; of 33 included studies, only one trial reported linkage adequately. Reporting of data validation and adjudication procedures, which is necessary to assess possible misclassification bias, was also not adequately reported in most trials. Another consistent gap related to implications of using administrative data, which is important for contextualizing trial results and understanding potential limitations of using administrative data in the trial. Finally, sources of funding for the administrative database used were rarely reported. Separate studies were conducted to evaluate reporting in trials conducted using electronic health records [19] and registries [20]. Similar trends were observed in those studies. In all trial types, items related to methodological considerations in using routinely collected data in trials, which were new CONSORT-ROUTINE items, were not adequately reported in most trials.

Our review has limitations that must be taken into account. First, our scoping review was able to capture only a sample of RCTs conducted using administrative databases rather than all trials that have been conducted using administrative databases. This was in part because of the lack of accepted specific Medical Subject Headings to identify RCTs conducted using administrative databases. In combination with our inclusion criteria on what constituted an RCT conducted using an administrative database, this led to a relatively small sample of only 33 RCTs. It is possible that this approach could have influenced the representativeness of the trials we included. For instance, we searched for trials based on their reporting of use of administrative data in the title or abstract; thus, it follows that this item would almost always be reported in our sample of trials ("Modified – Administrative database use and name in the abstract" and "Modified – Description of trial design"). Second, we did not extend our assessment to include study protocols for included trials. Some authors may have included additional study details within the protocol. However, the CONSORT extension checklist is a minimum set of standards that should be adequately reported in reports of trial outcomes, irrespective of having been previously published in a protocol or in a primary trial publication in the case of secondary reports.

5. Conclusion

In summary, this study was the first to assess the completeness and transparency of reporting of RCTs conducted using administrative databases against those elements now deemed to form a minimum reporting standard for such studies. Although we observed CONSORT 2010 criteria and items related to the application of the administrative database within the RCT to be largely adequately reported, we found a need for attention to more fulsome reporting of methodological conduct of these trials, mostly related to methodological aspects and implications of using administrative databases in RCTs. The new CONSORT-ROUTINE provides guidance to improve reporting of these types of trials. We recommend those who support, conduct, and report trials conducted using administrative databases to adhere to minimum reporting standards outlined in the newly developed CONSORT-ROUTINE, in order to ensure greater transparency and replicability and facilitate the use of trial results in healthcare decisions.

Availability of data and materials

Additional data beyond that reported in the main and supplementary materials can be requested from the corresponding author.

Author contributions

Mahrukh Imran: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Validation, Writing – original draft. **Kimberly McCord:** Conceptualization, Data curation, Investigation, Methodology, Validation, Writing – review & editing. **Stephen J. McCall:** Conceptualization, Data curation, Investigation, Methodology, Validation, Writing – review & editing. **Linda Kwakkenbos:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Supervision, Writing – review & editing. **Margaret Sampson:** Conceptualization, Methodology, Search, Writing – review & editing. **Ole Fröbert:** Conceptualization, Funding acquisition, Methodology, Writing – review & editing. **Chris Gale:** Conceptualization, Funding acquisition, Methodology, Writing – review & editing. **Lars G. Hemkens:** Conceptualization, Methodology, Supervision, Writing – review & editing. **Sinéad M. Langan:** Conceptualization, Methodology, Writing – review & editing. **David Moher:** Conceptualization, Methodology, Writing – review & editing. **Clare Relton:** Conceptualization, Funding acquisition, Methodology, Writing – review & editing. **Merrick Zwarenstein:** Conceptualization, Methodology, Writing – review & editing. **Edmund Juszczak:** Conceptualization, Funding acquisition, Methodology, Writing – review & editing. **Brett D. Thombs:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Supervision, Writing – review & editing.

Declaration of Competing Interests

All authors have completed the ICJME uniform disclosure form and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years

Acknowledgments

The development of CONSORT-ROUTINE and the present review were funded by grants from the Canadian

Institutes of Health Research (PI Thombs, #PJT-156172; PIs Thombs and Kwakkenbos, #PCS-161863) and from the United Kingdom National Institute of Health Research (NIHR) Clinical Trials Unit Support Funding (PI Juszczak, Co-PI Gale, supported salary of SM). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care. Dr. Langan was supported by a Wellcome Senior Clinical Fellowship in Science (205039/Z/16/Z). Dr. Moher is supported by a University Research Chair (uOttawa). Dr. Gale was supported by the United Kingdom Medical Research Council through a Clinician Scientist Fellowship. Dr. Thombs was supported by a Tier 1 Canada Research Chair.

Appendix 1. Electronic search strategies

Searches were run in both MEDLINE and Cochrane Methodology Register simultaneously. As an example, in the registries search, lines 1–11 are the MEDLINE search and lines 12–15 are tailored for the Cochrane Methodology Register. The final lines of each search isolate the records from each database, combine them so duplicate records can be removed, then isolate the remaining records so they can be downloaded and imported into Reference Manager using customized import filters.

Searches for RCTs conducted using Administrative Databases

1 randomized controlled trial.pt.
2 controlled clinical trial.pt.
3 randomi?ed.ab.
4 placebo.ab.
5 randomly.ab.
6 clinical trials as topic.sh.
7 trial.ti.
8 or/1-7
9 exp animals/ not humans.sh.
10 8 not 9
11 administrative data*.ab,kf,ti.
12 healthcare data*.ab,kf,ti.
13 health care data*.ab,kf,ti.
14 or/11-13
15 10 and 14
16 (administrative adj5 data*).ti,ab,kw.
17 health care data*.ti,ab,kw.
18 healthcare data*.ti,ab,kw.
19 or/16-18
20 (random* or RCT).ti,ab,kw.
21 19 and 20
22 limit 15 to yr="2007 - 2018"
23 22 use medall
24 limit 21 to yr="2007 - 2018"
25 22 use clcmr

Appendix 2. Inclusion/Exclusion criteria (Title and Abstract)

Exclude: not an RCT using administrative data. If it is clear from the title and abstract that the study is not an RCT using administrative data or is a publication on methods or reporting of RCTs using administrative data, it will be excluded. If it is clear from the title and abstract that the study only reports (1) issues related to methods or reporting of RCTs conducted using administrative data, or (2) is a protocol from a RCT conducted using administrative data, it is excluded. If the RCT involves non-human subjects, it is excluded. Only RCTs that use administrative data for conducting the trial, including activities such as identifying eligible participants for the trial or as an intervention or collecting trial outcomes, are eligible.

Include: the administrative database is used for identifying eligible participants. If it is clear from the title and abstract that the publication describes a trial in which the administrative database was used to identify eligible trial participants, it will be included.

Include: the administrative database is used to ascertain health outcomes. If it is clear from the title and abstract that the publication describes a trial that uses administrative data to ascertain health outcomes, as trial endpoints, it will be included.

Inclusion/Exclusion criteria (Full-text)

Exclude: not an RCT using administrative data. If the study is not an RCT using administrative data or is a publication on methods or reporting of RCTs using administrative data, it will be excluded. If the publication only reports (1) issues related to methods or reporting of RCTs conducted using administrative data, or (2) a protocol from a RCT conducted using administrative data, it is excluded. If the RCT involves non-human subjects, it is excluded. Only RCTs that use administrative data for conducting the trial, including activities such as identifying eligible participants for the trial or as an intervention or collecting trial outcomes, are eligible.

Include: the administrative database is used for identifying eligible participants. If the publication describes a trial in which the administrative database was used to identify eligible trial participants, it will be included.

Include: the administrative database is used to ascertain health outcomes. If the publication describes a trial that uses administrative data to ascertain health outcomes, as trial endpoints, it will be included.

Appendix 3. Coding manual for completeness and transparency of reporting

Please cite this article as: M. Imran et al., Reporting transparency and completeness in trials: Paper 3 – trials conducted using administrative databases do not adequately report elements related to use of databases, Journal of Clinical Epidemiology, https://doi.org/10.1016/j.jclinepi.2021.09.010

ORIGINAL CONSORT Item			CONSORT-ROUTINE	Adequately reported	Partially reported	Inadequately or Not reported	Not applicable
Title and abstract							
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts)		Did the authors clearly describe a (1) structured summary of (2) trial design, (3) methods, (4) results, and (5) conclusions.	Did the authors only report one, two, three or four element(s) of this item and not all five elements of the item?	Did the authors not describe a structured summary of trial design, methods, results and conclusions?	
			Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts). **Specify that a cohort or routinely collected data were used to conduct the trial and, if applicable, provide the name of the cohort or routinely collected database(s) (Modified)**	Did the authors specify that a routinely collected database(s) was used to conduct the trial? (Sufficient to detail that an "administrative database was used").	Did the authors describe methods that would typically require a routinely collected database for components of the trials but **not** specify they used a routinely collected database(s)?	Did the authors **not** specify that a routinely collected database(s) was used to conduct the trial?	
Methods							
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio		Did the authors clearly describe the trial design including allocation ratio?	All other cases, where applicable.	Did the authors **not** describe the trial design including allocation ratio?	
			Description of trial design (such as parallel, factorial) including allocation ratio, **that a cohort or routinely collected database(s) was used to conduct the trial (such as electronic health record, registry) and how the data were used within the trial (such as identification of eligible trial participants, trial outcomes) (Modified)**	Did the authors clearly mention that **(1)** a routinely collected database(s) was used within the trial and **(2)** how the data were used within the trial (i.e. identification of participants, outcome measurement, other)?	Did the authors only report one element of this item and **not** both elements of the item?	Did the authors **not** state that a routinely collected database(s) was used within the trial **and** not describe how the data were used within the trial (i.e. identification of participants, outcome measurement, other)?	
Cohort or routinely collected database							
	ROUTINE-1		Name, if applicable, and description of the cohort or routinely collected database(s) used to conduct the trial, including information on the setting (such as primary care), locations, and dates, (such as periods of recruitment, follow-up, and data collection) (New)	Did the authors clearly **(1)** name and**(2)** describe the routinely collected database(s) and **(3)** provide information on the setting, locations, and relevant dates (e.g. periods of recruitment, follow-up, and data collection)?	Did the authors only report one or two element(s) of this item and **not** all three element**s** of the item?	Did the authors **not** name **and** describe the routinely collected database(s) **and not** provide information on the setting, locations, **and** relevant dates (e.g. periods of recruitment, follow-up, and data collection)?	

(*continued on next page*)

Please cite this article as: M. Imran et al., Reporting transparency and completeness in trials: Paper 3 – trials conducted using administrative databases do not adequately report elements related to use of databases, Journal of Clinical Epidemiology, https://doi.org/10.1016/j.jclinepi.2021.09.010

(*continued*)

	ROUTINE-2		Eligibility criteria for participants in the cohort or routinely collected database(s) (New)	Did the authors clearly describe eligibility criteria for the routinely collected database(s)?	All other cases, where applicable.	Did the authors **not** describe all eligibility criteria for the routinely collected database(s)?	
	ROUTINE-3		State whether the study included person-level, institutional-level, or other data linkage across two or more databases and, if so, linkage techniques and methods used to evaluate completeness and accuracy of linkage (New)	Did the authors clearly state whether the study included **(1)** person-level, institutional-level, or other data linkage across two or more databases **and (2)** the methods of linkage **and (3)** methods used to evaluate completeness and accuracy of linkage?	Did the authors only report one element of this item and **not** all three elements of the item?	Did the authors **not** state whether the study included person-level, institutional-level, or other data linkage across two or more databases **and** not state the methods of linkage **and** methods used to evaluate completeness and accuracy of linkage?	
Trial participants	4a	Eligibility criteria for participants		Did the authors clearly describe the eligibility for the trial participants?	All other cases, where applicable.	Did the authors **not** describe all eligibility criteria for the trial participants?	
			Eligibility criteria for trial participants, **including information on how to access the list of codes and algorithms used to identify eligible participants, including methods used to assess accuracy and completeness, if applicable (Modified)**	Did the authors provide information on **(1)** how to access the lists of codes and algorithms used to identify participants, including **(2)** methods used to assess accuracy and completeness, if applicable?	Did the authors only report one element of this item and **not** both elements of the item?	Did the authors **not** provide information on how to access the lists of codes and algorithms used to identify participants, **and not** provide the methods used to assess accuracy and completeness?	The trial did not use routinely collected data to identify participants
	ROUTINE-4		Describe whether and how consent was obtained (New)	Did the authors describe clearly whether and how consent was obtained?	All other cases, where applicable.	Did the authors **not** describe whether and how consent was obtained?	
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed		Did the authors clearly define the pre-specified primary and secondary outcome measures, including how and when they were assessed?	Did the authors **only** define the pre-specified primary and secondary outcome measures but **not** how and when they were assessed or did they describe how and when outcomes were assessed but not the measures?	Did the authors **not** define the pre-specified primary and secondary outcome measures **and not** define how and when they were assessed?	

(*continued on next page*)

(*continued*)

			Completely defined pre-specified primary and secondary outcome measures, including how and when they were ascertained and **the cohort or routinely collected database(s) used to ascertain each outcome (Modified)**	Did the authors clearly describe the routinely collected database(s) used to ascertain each outcome?	All other cases, where applicable.	Did the authors **not** describe the routinely collected database(s) used to ascertain each outcome?	The trial did not use routinely collected data to ascertain the outcome
	ROUTINE-5		Information on how to access the list of codes and algorithms used to define or derive the outcomes from the cohort or routinely collected database(s) used to conduct the trial, including methods used to assess accuracy and completeness, if applicable (New)	Did the authors clearly **(1)** describe information on how to access the list of codes and algorithms used to define or derive the outcomes from the routinely collected database(s), **(2)** including methods used to assess accuracy and completeness?	Did the authors only report one element of this item and **not** both elements of the item?	Did the authors **not** describe information on how to access the list of codes and algorithms used to define or derive the outcomes from the routinely collected database(s), **and not** describe the methods used to assess accuracy and completeness?	The trial did not use routinely collected data to ascertain the outcome
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Mechanism used to implement the random allocation sequence **(such as embedding an automated randomiser within the cohort or routinely collected database(s))**, describing any steps taken to conceal the sequence until interventions were assigned (Modified)	Did the authors clearly describe the mechanism used to implement the random allocation sequence (such as embedding an automated randomiser within the cohort or routinely collected database(s)), describing any steps taken to conceal the sequence until interventions were assigned?	All other cases, where applicable	Did the authors **not** describe the mechanism used to implement the random allocation sequence (such as embedding an automated randomiser within the cohort or routinely collected database(s)), describing any steps taken to conceal the sequence until interventions were assigned?	
Results							
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome		Did the authors define clearly for each group, **(1)** the number of participants who were randomly assigned, **(2)** received intended treatment and **(3)** were analysed for the primary outcome?	Did the authors only report one or two elements of this item and **not** all three elements of the item **or** only presented this information for one group?	Did the authors **not** describe clearly for each group, the number of participants who were randomly assigned, **and not** received intended treatment **and not** were analysed for the primary outcome?	

(*continued on next page*)

(*continued*)

			For each group, the number of participants in the cohort or routinely collected database(s) used to conduct the trial and the numbers screened for eligibility, randomly assigned, offered and accepted interventions (e.g., cohort multiple RCTs), received intended treatment, and analysed for the primary outcome (Modified)	Did the authors clearly define, for each group, the number of participants in the routinely collected database(s) used to conduct the trial and the numbers screened for eligibility, randomly assigned, received intended treatment, and analysed for the primary outcome?	Did the authors only report **some, but not all**, elements of this item?	Did the authors **not** define, for each group, the number of participants in the routinely collected database(s) used to conduct the trial **and not** define the numbers screened for eligibility, randomly assigned, received intended treatment, and analysed for the primary outcome
Discussion						
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		Did the authors clearly provide an interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence?	All other cases, where applicable	Did the authors **not** provide an interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence?
			Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence, **including the implications of using data that were not collected to answer the trial research questions** (Modified)	Did the authors (1) clearly provide an interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence **and** (2) describe the implications of using data that were not collected to answer the trial research questions?	Did the authors (1) clearly provide an interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence **or** (2) describe the implications of using data that were not collected to answer the trial research questions – but not both?	Did the authors **not** provide an interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence **and not** describe the implications of using data that were not collected to answer the trial research questions?
Other information						
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders		Did the authors clearly describe the sources of funding and the role of funders?	All other cases, where applicable	Did the authors **not** describe the sources of funding and other support for the trial **and** the role of the funders?
			Sources of funding and other support for both the trial **and the cohort or routinely collected database(s)**, role of funders (Modified)	Did the authors clearly describe the sources of funding for the database(s) and trial and the role of the funder of the trial?	Did the authors only report **some, but not all**, elements of this item?	Did the authors **not** describe the sources of funding for routinely collected database(s) and trial and **not** describe the role of the funder of the trial?

Appendix 4. References of 33 included trials

[1] Ahmed S, Ernst P, Bartlett SJ, Valois MF, Zaihra T, Pare G, et al. The Effectiveness of Web-Based Asthma Self-Management System, My Asthma Portal (MAP): A Pilot Randomized Controlled Trial. Journal of Medical Internet Research. 2016;18:e313.

[2] Armstrong CD, Taljaard M, Hogg W, Mark AE, Liddy C. Practice facilitation for improving cardiovascular care: secondary evaluation of a stepped wedge cluster randomized controlled trial using population-based administrative data. Trials. 2016;17:434.

[3] Bell JF, Krupski A, Joesch JM, West II, Atkins DC, Court B, et al. A randomized controlled trial of intensive care management for disabled Medicaid beneficiaries with high health care costs. Health Services Research. 2015;50:663.

[4] Belsher BE, Jaycox LH, Freed MC, Evatt DP, Liu X, Novak LA, et al. Mental Health Utilization Patterns During a Stepped, Collaborative Care Effectiveness Trial for PTSD and Depression in the Military Health System. Medical Care. 2016;54:706.

[5] Cushman WC, Davis BR, Pressel SL, Cutler JA, Einhorn PT, Ford CE, et al. Mortality and morbidity during and after the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial. Journal of Clinical Hypertension. 2012;14:20.

[6] Desai S, Mahal A, Sinha T, Schellenberg J, Cousens S. The effect of community health worker-led education on women's health and treatment-seeking: A cluster randomised trial and nested process evaluation in Gujarat, India. Journal of Global Health. 2017;7:020404.

[7] Englander H, Michaels L, Chan B, Kansagara D. The care transitions innovation (C-TraIn) for socioeconomically disadvantaged adults: results of a cluster randomized controlled trial. Journal of General Internal Medicine. 2014;29:1460.

[8] French SD, McKenzie JE, Connor DAO, Grimshaw JM, Mortimer D, Francis JJ, et al. Evaluation of a theory-informed implementation intervention for the management of acute low back pain in general medical practice: the IMPLEMENT cluster randomised trial. PLoS ONE [Electronic Resource]. 2013;8:e65471.

[9] Goldman LE, Sarkar U, Kessell E, Guzman D, Schneidermann M, Pierluissi E, et al. Support from hospital to home for elders: a randomized trial. Annals of Internal Medicine. 2014;161:472.

[10] Harron K, Mok Q, Dwan K, Ridyard CH, Moitt T, Millar M, et al. CATheter Infections in CHildren (CATCH): a randomised controlled trial and economic evaluation comparing impregnated and standard central venous catheters in children. Health Technology Assessment (Winchester, England). 2016;20:vii.

[11] Haywood LJ, Davis BR, Piller LB, Cushman WC, Cutler JA, Ford CE, et al. Influence of Prevalent and Incident Atrial Fibrillation on Post-Trial Major Events in ALLHAT. Journal of the National Medical Association. 2017;109:172.

[12] Jacobs F, Easterbrooks MA, Goldberg J, Mistry J, Bumgarner E, Raskin M, et al. Improving Adolescent Parenting: Results From a Randomized Controlled Trial of a Home Visiting Program for Young Families. American Journal of Public Health. 2016;106:342.

[13] Keall MD, Pierse N, Howden-Chapman P, Cunningham C, Cunningham M, Guria J, et al. Home modifications to reduce injuries from falls in the home injury prevention intervention (HIPI) study: a cluster-randomised controlled trial. Lancet. 2015;385:231.

[14] Levine DA, Funkhouser EM, Houston TK, Gerald JK, Johnson-Roe N, Allison JJ, et al. Improving care after myocardial infarction using a 2-year internet-delivered intervention: the Department of Veterans Affairs myocardial infarction-plus cluster-randomized trial. Archives of Internal Medicine. 2011;171:1910.

[15] Murdoch M, Simon AB, Polusny MA, Bangerter AK, Grill JP, Noorbaloochi S, et al. Impact of different privacy conditions and incentives on survey response rate, participant representativeness, and disclosure of sensitive information: a randomized controlled trial. BMC Medical Research Methodology. 2014;14:90.

[16] Palacio AM, Uribe C, Hazel-Fernandez L, Li H, Tamariz LJ, Garay SD, et al. Can phone-based motivational interviewing improve medication adherence to antiplatelet medications after a coronary stent among racial minorities? A randomized trial. Journal of General Internal Medicine. 2015;30:469.

[17] Patel R, Powell JT, Sweeting MJ, Epstein DM, Barrett JK, Greenhalgh RM. The UK EndoVascular Aneurysm Repair (EVAR) randomised controlled trials: long-term follow-up and cost-effectiveness analysis. Health technology assessment (Winchester, England). 2018;22:1.

[18] Piller LB, Baraniuk S, Simpson LM, Cushman WC, Massie BM, Einhorn PT, et al. Long-term follow-up of participants with heart failure in the antihypertensive and lipid-lowering treatment to prevent heart attack trial (ALLHAT). Circulation. 2011;124:1811.

[19] Piller LB, Simpson LM, Baraniuk S, Habib GB, Rahman M, Basile JN, et al. Characteristics and long-term follow-up of participants with peripheral arterial disease during ALLHAT. Journal of General Internal Medicine. 2014;29:1475.

[20] Rappaport EB, Daskalakis C, Sendecki JA. Using routinely collected growth data to assess a school-based obesity prevention strategy. International Journal of Obesity. 2013;37:79.

[21] Rosen CS, Azevedo KJ, Tiet QQ, Greene CJ, Wood AE, Calhoun P, et al. An RCT of Effects of Telephone Care Management on Treatment Adherence and Clinical Outcomes Among Veterans With PTSD. Psychiatric Services. 2017;68:151.

[22] Rosen CS, Tiet QQ, Harris AH, Julian TF, McKay JR, Moore WM, et al. Telephone monitoring and support

Please cite this article as: M. Imran et al., Reporting transparency and completeness in trials: Paper 3 – trials conducted using administrative databases do not adequately report elements related to use of databases, Journal of Clinical Epidemiology, https://doi.org/10.1016/j.jclinepi.2021.09.010

after discharge from residential PTSD treatment: a randomized controlled trial. Psychiatric Services. 2013;64:13.

[23] Sandner M, Cornelissen T, Jungmann T, Herrmann P. Evaluating the effects of a targeted home visiting program on maternal and child health outcomes. Journal of Health Economics. 2018;58:269.

[24] Schmidt-Mende K, Andersen M, Wettermark B, Hasselstrom J. Educational intervention on medication reviews aiming to reduce acute healthcare consumption in elderly patients with potentially inappropriate medicines-A pragmatic open-label cluster-randomized controlled trial in primary care. Pharmacoepidemiology & Drug Safety. 2017;26:1347.

[25] Shah BR, Bhattacharyya O, Yu CH, Mamdani MM, Parsons JA, Straus SE, et al. Effect of an educational toolkit on quality of care: a pragmatic cluster randomized trial. PLoS Medicine / Public Library of Science. 2014;11:e1001588.

[26] Siddique HH, Olson RH, Parenti CM, Rector TS, Caldwell M, Dewan NA, et al. Randomized trial of pragmatic education for low-risk COPD patients: impact on hospitalizations and emergency department visits. International Journal of Copd. 2012;7:719.

[27] Somers JM, Patterson ML, Moniruzzaman A, Currie L, Rezansoff SN, Palepu A, et al. Vancouver At Home: pragmatic randomized trials investigating Housing First for homeless and mentally ill adults. Trials [Electronic Resource]. 2013;14:365.

[28] Steventon A, Bardsley M, Billings J, Dixon J, Doll H, Beynon M, et al. Effect of telecare on use of health and social care services: findings from the Whole Systems Demonstrator cluster randomised trial. Age & Ageing. 2013;42:501.

[29] Steventon A, Bardsley M, Billings J, Dixon J, Doll H, Hirani S, et al. Effect of telehealth on use of secondary care and mortality: findings from the Whole System Demonstrator cluster randomised trial. BMJ (Clinical research ed). 2012;344:e3874.

[30] Tamblyn R, Eguale T, Buckeridge DL, Huang A, Hanley J, Reidel K, et al. The effectiveness of a new generation of computerized drug alerts in reducing the risk of injury from drug side effects: a cluster randomized trial. Journal of the American Medical Informatics Association. 2012;19:635.

[31] Vandivere S, Malm KE, Allen TJ, Williams SC, McKlindon A. A Randomized Controlled Trial of Family Finding: A Relative Search and Engagement Intervention for Youth Lingering in Foster Care. Evaluation Review. 2017;41:542.

[32] Wagner EH, Ludman EJ, Bowles EJi, Penfold R, Reid RJ, Rutter CM, et al. Nurse navigators in early cancer care: a randomized, controlled trial. Journal of Clinical Oncology. 2014;32:12.

[33] Zeiger RS, Schatz M, Li Q, Solari PG, Zazzali JL, Chen W. Real-time asthma outreach reduces excessive short-acting beta2-agonist use: a randomized study. The Journal of Allergy & Clinical Immunology in Practice. 2014;2:445.

Appendix 5. Table. Characteristics of trials conducted using administrative databases by cluster versus individually randomized trials

Please cite this article as: M. Imran et al., Reporting transparency and completeness in trials: Paper 3 – trials conducted using administrative databases do not adequately report elements related to use of databases, Journal of Clinical Epidemiology, https://doi.org/10.1016/j.jclinepi.2021.09.010

	Number (%) of cluster randomized trials (n = 13)	Number (%) of individually randomized trials (n = 20)	Total (%) (n = 33)
Publication type			
Primary	11 (85%)	14 (70%)	25 (76%)
Secondary	2 (15%)	6 (30%)	8 (24%)
Use of administrative data in trial			
Identification of patients	0 (0%)	1 (5%)	1 (3%)
Outcome ascertainment	10 (77%)	15 (75%)	25 (76%)
Both identification of patients and outcome ascertainment	3 (23%)	4 (20%)	7 (21%)
Administrative data used for primary outcome			
Yes	8 (62%)	14 (70%)	22 (67%)
No	5 (38%)	5 (25%)	10 (30%)
Unclear	0 (0%)	1 (5%)	1 (3%)
Setting			
Inpatient	3 (23%)	8 (40%)	11 (33%)
Primary care	7 (54%)	3 (15%)	10 (30%)
Community medicine	2 (15%)	3 (15%)	5 (15%)
Outpatient	0 (0%)	3 (15%)	3 (9%)
Other[1]	1 (8%)	3 (15%)	4 (12%)
Country			
USA	4 (31%)	14 (70%)	18 (55%)
Canada	3 (23%)	3 (15%)	6 (18%)
UK	2 (15%)	2 (10%)	4 (12%)
Europe	1 (8%)	1 (5%)	2 (6%)
Australia	1 (8%)	0 (0%)	1 (3%)
India	1 (8%)	0 (0%)	1 (3%)
New Zealand	1 (8%)	0 (0%)	1 (3%)
Disease type			
General medicine/health	7 (54%)	5 (25%)	12 (36%)
Cardiovascular disease	3 (23%)	6 (30%)	9 (27%)
Mental health	0 (0%)	3 (15%)	3 (9%)
Respiratory	0 (0%)	3 (15%)	3 (9%)
Other[2]	3 (23%)	3 (15%)	6 (18%)
Intervention			
Educational	3 (23%)	7 (35%)	10 (30%)
Multicomponent	5 (38%)	2 (10%)	7 (21%)
Drug	0 (0%)	4 (20%)	4 (12%)
Guideline/reminder-based	2 (2%)	0 (0%)	2 (6%)
Other[3]	3 (23%)	7 (35%)	10 (30%)
Comparator			
Usual care	12 (92%)	13 (65%)	25 (76%)
Active comparator	1 (8%)	7 (35%)	8 (24%)
Primary outcome			
Mortality	0 (0%)	5 (25%)	5 (15%)
Hospitalization	3 (23%)	2 (10%)	5 (15%)

(*continued on next page*)

Please cite this article as: M. Imran et al., Reporting transparency and completeness in trials: Paper 3 – trials conducted using administrative databases do not adequately report elements related to use of databases, Journal of Clinical Epidemiology, https://doi.org/10.1016/j.jclinepi.2021.09.010

(*continued*)

Surrogate	3 (23%)	1 (5%)	4 (12%)
Self-reported	1 (8%)	2 (10%)	3 (9%)
Other[4]	5 (38%)	11 (55%)	16 (48%)
Sample size			
Clusters (Median and IQR)	101 [373–221]		101 [73–221]
Participants (Median and IQR)	119,910 [86,998–526,850]	32,804 [32,804–33,081]	43,721 [32,942–103,453]

[1] Residential and multiple.
[2] Diabetes, cancer, potentially inapproproate medicines, drug side effects, infection, disability, homelessness.
[3] Telephone/web-based care, Family Finding program, referral, housing, health care provider support, surgical.
[4] Insurance claims, uptake of treatment, disease occurence, no primary outcome, adherence, risk of injury, multiple/composite outcomes and rate of injury.

Appendix 6. Completeness and transparency of reporting for each item for RCTs conducted using administrative data by Primary/Secondary publication type (only includes original, modified and new items)

Please cite this article as: M. Imran et al., Reporting transparency and completeness in trials: Paper 3 – trials conducted using administrative databases do not adequately report elements related to use of databases, Journal of Clinical Epidemiology, https://doi.org/10.1016/j.jclinepi.2021.09.010

CONSORT 2010 Item			CONSORT-ROUTINE	Primary Publications, N=25			
				Adequately reported N (%)	Partially reported N (%)	Inadequately or Not reported N (%)	Not applicable N (%)
Title and abstract							
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts).		22 (80%)	3 (12%)	0 (0%)	-
			Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts). **Specify that a cohort or routinely collected data were used to conduct the trial and, if applicable, provide the name of the cohort or routinely collected database(s) (Modified)**	22 (80%)	3 (12%)	0 (0%)	-
Methods							
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio		8 (32%)	6 (24%)	11 (44%)	-
			Description of trial design (such as parallel, factorial) including allocation ratio, **that a cohort or routinely collected database(s) was used to conduct the trial (such as electronic health record, registry) and how the data were used within the trial (such as identification of eligible trial participants, trial outcomes) (Modified)**	20 (80%)	5 (20%)	0 (0%)	-
Cohort or routinely collected database	ROUTINE-1		Name, if applicable, and description of the cohort or routinely collected database(s) used to conduct the trial, including information on the setting (such as primary care), locations, and dates, (such as periods of recruitment, follow-up, and data collection) (New)	2 (8%)	20 (80%)	3 (12%)	-
	ROUTINE-2		Eligibility criteria for participants in the cohort or routinely collected database(s) (New)	1 (4%)	5 (20%)	19 (76%)	-
	ROUTINE-3		State whether the study included person-level, institutional-level, or other data linkage across two or more databases and, if so, linkage techniques and methods used to evaluate completeness and accuracy of linkage (New)	1 (4%)	8 (32%)	16 (64%)	-

(*continued on next page*)

Please cite this article as: M. Imran et al., Reporting transparency and completeness in trials: Paper 3 – trials conducted using administrative databases do not adequately report elements related to use of databases, Journal of Clinical Epidemiology, https://doi.org/10.1016/j.jclinepi.2021.09.010

(*continued*)

Trials participants	4a	Eligibility criteria for participants		20 (80%)	4 (16%)	1 (4%)	-
			Eligibility criteria for trial participants, **including information on how to access the list of codes and algorithms used to identify eligible participants, information on accuracy and completeness of data used to ascertain eligibility, and methods used to validate accuracy and completeness (e.g., monitoring, adjudication), if applicable (Modified)**	0 (0%)	4 (16%)	0 (0%)	21 (84%)
	ROUTINE-4		Describe whether and how consent was obtained (New)	19 (76%)	1 (4%)	5 (20%)	-
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed		23 (92%)	2 (8%)	0 (0%)	-
			Completely defined pre-specified primary and secondary outcome measures, including how and when they were ascertained and **the cohort or routinely collected database(s) used to ascertain each outcome (Modified)**	21 (84%)	4 (16%)	0 (0%)	0 (0%)
	ROUTINE-5		Information on how to access the list of codes and algorithms used to define or derive the outcomes from the cohort or routinely collected database(s) used to conduct the trial, information on accuracy and completeness of outcome variables, and methods used to validate accuracy and completeness (e.g., monitoring, adjudication), if applicable. (New)	0 (0%)	2 (8%)	22 (88%)	1 (4%)
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Mechanism used to implement the random allocation sequence **(such as embedding an automated randomiser within the cohort or routinely collected database(s)),** describing any steps taken to conceal the sequence until interventions were assigned (Modified)	8 (32%)	2 (8%)	15 (60%)	-
Results							
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome		17 (68%)	7 (28%)	1 (4%)	-

(*continued on next page*)

Please cite this article as: M. Imran et al., Reporting transparency and completeness in trials: Paper 3 – trials conducted using administrative databases do not adequately report elements related to use of databases, Journal of Clinical Epidemiology, https://doi.org/10.1016/j.jclinepi.2021.09.010

(*continued*)

			For each group, the number of participants **in the cohort or routinely collected database(s) used to conduct the trial and the numbers screened for eligibility,** randomly assigned, **offered and accepted interventions (e.g., cohort multiple RCTs),** received intended treatment, and analysed for the primary outcome (Modified)	1 (4%)	2 (8%)	1 (4%)	21 (84%)
Discussion							
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		24 (96%)	1 (4%)	0 (0%)	-
			Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence, **including the implications of using data that were not collected to answer the trial research questions (Modified)**	5 (20%)	1 (4%)	19 (76%)	-
Other information							
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders		12 (48%)	12 (48%)	1 (4%)	-
			Sources of funding and other support **for** both the trial and **the cohort or routinely collected database(s),** role of funders (Modified)	2 (8%)	13 (52%)	40 (0%)	-
				Secondary Publications, N=8			
Title and abstract							
	1b	Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts).		7 (88%)	1 (12%)	0 (0%)	-
			Structured summary of trial design, methods, results, and conclusions (for specific guidance see CONSORT for abstracts). **Specify that a cohort or routinely collected data were used to conduct the trial and, if applicable, provide the name of the cohort or routinely collected database(s) (Modified)**	8 (100%)	0 (0%)	0 (0%)	-
Methods							
Trial design	3a	Description of trial design (such as parallel, factorial) including allocation ratio		3 (38%)	3 (38%)	2 (25%)	-

(*continued on next page*)

(*continued*)

			Description of trial design (such as parallel, factorial) including allocation ratio, **that a cohort or routinely collected database(s) was used to conduct the trial (such as electronic health record, registry) and how the data were used within the trial (such as identification of eligible trial participants, trial outcomes) (Modified)**	7 (88%)	1 (13%)	0 (0%)	-
Cohort or routinely collected database	ROUTINE-1		Name, if applicable, and description of the cohort or routinely collected database(s) used to conduct the trial, including information on the setting (such as primary care), locations, and dates, (such as periods of recruitment, follow-up, and data collection) (New)	1 (13%)	7 (88%)	0 (0%)	-
	ROUTINE-2		Eligibility criteria for participants in the cohort or routinely collected database(s) (New)	1 (13%)	2 (25%)	5 (63%)	-
	ROUTINE-3		State whether the study included person-level, institutional-level, or other data linkage across two or more databases and, if so, linkage techniques and methods used to evaluate completeness and accuracy of linkage (New)	0 (0%)	3 (38%)	5 (63%)	-
Trials participants	4a	Eligibility criteria for participants		8 (100%)	0 (0%)	0 (0%)	-
			Eligibility criteria for trial participants, **including information on how to access the list of codes and algorithms used to identify eligible participants, information on accuracy and completeness of data used to ascertain eligibility, and methods used to validate accuracy and completeness (e.g., monitoring, adjudication), if applicable (Modified)**	0 (0%)	1 (13%)	1 (13%)	6 (75%)
	ROUTINE-4		Describe whether and how consent was obtained (New)	7 (88%)	0 (0%)	1 (12%)	-
Outcomes	6a	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed		8 (100%)	0 (0%)	0 (0%)	-
			Completely defined pre-specified primary and secondary outcome measures, including how and when they were ascertained and **the cohort or routinely collected database(s) used to ascertain each outcome (Modified)**	8 (100%)	0 (0%)	0 (0%)	0 (0%)

(*continued on next page*)

(*continued*)

	ROUTINE-5		Information on how to access the list of codes and algorithms used to define or derive the outcomes from the cohort or routinely collected database(s) used to conduct the trial, information on accuracy and completeness of outcome variables, and methods used to validate accuracy and completeness (e.g., monitoring, adjudication), if applicable. (New)	0 (0%)	3 (38%)	5 (63%)	-
Allocation concealment mechanism	9	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	Mechanism used to implement the random allocation sequence **(such as embedding an automated randomiser within the cohort or routinely collected database(s)),** describing any steps taken to conceal the sequence until interventions were assigned (Modified)	1 (13%)	1 (13%)	6 (75%)	-
Results							
Participant flow (a diagram is strongly recommended)	13a	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome		5 (63%)	2 (25%)	1 (13%)	-
			For each group, the number of participants **in the cohort or routinely collected database(s) used to conduct the trial and the numbers screened for eligibility,** randomly assigned, **offered and accepted interventions (e.g., cohort multiple RCTs),** received intended treatment, and analysed for the primary outcome (Modified)	0 (0%)	3 (38%)	0 (0%)	5 (63%)
Discussion							
Interpretation	22	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence		8 (100%)	0 (0%)	0 (0%)	-
			Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence, **including the implications of using data that were not collected to answer the trial research questions (Modified)**	2 (25%)	0 (0%)	6 (75%)	-
Other information							
Funding	25	Sources of funding and other support (such as supply of drugs), role of funders		7 (88%)	1 (13%)	0 (0%)	-
			Sources of funding and other support **for** both the trial and **the cohort or routinely collected database(s), role of funders (Modified)**	0 (0%)	7 (88%)	1 (13%)	-

REFERENCES

[1] Anderson GL, Burns CJ, Larsen J, Shaw PA. Use of administrative data to increase the practicality of clinical trials: insights from the Women's Health Initiative. Clin Trials 2016;13:519–26.
[2] Mazzali C, Duca P. Use of administrative data in healthcare research. Internal Emerg Med 2015;10:517–24.
[3] Hashimoto RE, Brodt ED, Skelly AC, Dettori JR. Administrative database studies: goldmine or goose chase? Evidence Based Spine–Care J 2014;5:074–6.
[4] Cadarette SM, Wong L. An introduction to health care administrative data. Can J Hosp Pharm 2015;68:232.
[5] Mc Cord KA, Salman RA-S, Treweek S, Gardner H, Strech D, Whiteley W, et al. Routinely collected data for randomized trials: promises, barriers, and implications. Trials 2018;19:29.
[6] Khan A, Ramsey K, Ballard C, Armstrong E, Burchill LJ, Menashe V, et al. Limited accuracy of administrative data for the identification and classification of adult congenital heart disease. J Am Heart Assoc 2018;7:e007378.
[7] Peabody JW, Luck J, Jain S, Bertenthal D, Glassman P. Assessing the accuracy of administrative data in health information systems. Med Care 2004:1066–72.
[8] Harron K, Dibben C, Boyd J, Hjern A, Azimaee M, Barreto ML, et al. Challenges in administrative data linkage for research. Big Data Soc 2017;4 2053951717745678.
[9] Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. Trials 2010;11:32.
[10] Juszczak E, Altman DG, Hopewell S, Schulz K. Reporting of multi-arm parallel-group randomized trials: extension of the CONSORT 2010 statement. JAMA 2019;321:1610–20.
[11] Campbell MK, Piaggio G, Elbourne DR, Altman DG. Consort 2010 statement: extension to cluster randomised trials. BMJ 2012;345:e5661.
[12] Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ 2016;355:i5239.
[13] Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ 2008;337:a2390.
[14] Kwakkenbos L, Imran M, McCall S, Mc Cord KA, Fröbert O, Hemkens LG, et al. CONSORT extension for the reporting of randomised controlled trials conducted using cohorts and routinely collected data: checklist with explanation and elaboration. Under review.
[15] Kwakkenbos L, Imran M, McCord KA, Sampson M, Fröbert O, Gale C, et al. Protocol for a scoping review to support development of a CONSORT extension for randomised controlled trials using cohorts and routinely collected health data. BMJ Open 2018;8:e025266.
[16] McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol 2016;75:40–6.
[17] Distiller SR. This is a software program, so these aspects are not relevant. Evidence Partners. Ottawa, Canada. 2021.
[18] Khangura S, Konnyu K, Cushman R, Grimshaw J, Moher D. Evidence summaries: the evolution of a rapid review approach. Syst Rev 2012;1:10.
[19] McCall SJ, Imran M, Hemkens LG, Mc Cord K, Kwakkenbos L, Sampson M, et al. Reporting of randomised controlled trials conducted using electronic health records – room for improvement. Under review. 2021
[20] Mc Cord KA, Imran M, McCall SJ, Kwakkenbos L, Sampson M, Fröbert O, et al. Reporting of randomised trials using registries was mostly inadequate and hindered interpretation of results. Under review. 2021

Please cite this article as: M. Imran et al., Reporting transparency and completeness in trials: Paper 3 – trials conducted using administrative databases do not adequately report elements related to use of databases, Journal of Clinical Epidemiology, https://doi.org/10.1016/j.jclinepi.2021.09.010