

BMJ Open Systematic review protocol assessing the processes for linking clinical trial registries and their published results

Rabia Bashir, Adam G Dunn

To cite: Bashir R, Dunn AG. Systematic review protocol assessing the processes for linking clinical trial registries and their published results. *BMJ Open* 2016;**6**:e013048. doi:10.1136/bmjopen-2016-013048

► Prepublication history and additional material is available. To view please visit the journal (<http://dx.doi.org/10.1136/bmjopen-2016-013048>).

Received 15 June 2016
Revised 3 August 2016
Accepted 13 September 2016



CrossMark

Centre for Health Informatics,
Australian Institute of Health
Innovation, Macquarie
University, Sydney,
New South Wales, Australia

Correspondence to Rabia
Bashir; rabia.bashir@students.mq.edu.au

ABSTRACT

Introduction: Clinical trial registries are an important source of information for tracking clinical trials from their inception through to their reporting, and have been used to measure publication bias and outcome reporting bias. Our aim is to survey and quantify the processes that have been used to identify links between clinical trial registries and published trial reports in studies that rely on these links to evaluate the completeness and accuracy of trial reporting.

Methods and analysis: We will identify studies that describe a process for identifying the links between a trial registry included in the WHO International Clinical Trial Registry Platform and published trial results, and use those links to evaluate the completeness and accuracy of trial reporting. Information extracted from the studies will include the purpose and application domain of the study, registries used or searched, processes by which the links were identified, the study period and proportions for which links were found. We will summarise what is known about the number and availability of links between clinical trial registries and published results, and examine how automatic linking, inference and inquiry processes have been used to identify links since the introduction of trial registries.

Ethics and dissemination: The systematic review is focused on the analysis of secondary data and does not require ethics approval. The results of the systematic review will be used to inform standard processes used to identify links to and from clinical trial registries in studies that evaluate the completeness and accuracy of clinical trial reports, as well as systematic reviews. Our findings will be disseminated by publishing the systematic review in a peer-reviewed journal, and by engaging with stakeholders from clinical trial registries and bibliographic databases.

INTRODUCTION

Clinical trial registries were designed to provide information to researchers, clinicians and the public about trials that are underway or for which the results have not been reported.^{1 2} Since their introduction, their use has increased substantially following changes in requirements for journal publication and changes to the law in several

Strengths and limitations of this study

- This systematic review will quantify the processes used to link clinical trial registries to clinical trial results and determine how these may have changed since the introduction of clinical trial registries.
- The processes used to link clinical trial registries to published reports of clinical trials vary across studies that rely on those links to evaluate the accuracy and completeness of trial reports, and this systematic review will quantify these differences to inform the way this is performed in the future.
- By producing a baseline measurement of the availability of automatic links and the number of other links that must be identified through inquiry or inference, the systematic review will help determine the potential value of using clinical trial registries to augment current methods used to identify trials for systematic reviews of clinical interventions.
- As studies linking clinical trial registry data to published results are designed for a range of different purposes, the processes used to identify links are not always reported completely, making information extraction difficult.

countries,^{3–5} and a number of studies have examined publication bias^{6–17} and outcome reporting bias,^{18–27} using one or more of the registries.

Studies that use clinical trial registries to examine the completeness and accuracy of clinical trial reporting rely on being able to establish links between registries and reports of clinical trials. A proportion of those links can be accessed automatically,^{28 29} but the remainder must be determined by inference or inquiry. The manner in which these processes are used vary from study to study and are known to be time-consuming.^{9 26 27} It is not yet known whether differences in the way links are established by these processes have influenced the results of studies examining publication bias or outcome reporting bias. As a consequence, there is a current need to

survey the studies that have used these processes to identify links between clinical trial registries and their published results.

Clinical trial registries are sometimes used to identify trials for inclusion in systematic reviews.^{30–32} Systematic reviews benefit from clinical trial registries because they can be used to quantify reporting bias for an intervention or condition, can be used to assist in scheduling updates,^{33–36} also could be used as an external corpus in machine learning methods that automate or assist in searching and screening methods³⁷ and in some cases as a source of trial results that have not been published in peer-reviewed literature.^{31 32} By understanding the processes that have been used to establish links between trial registries and published trial reports, we may be able to provide guidance on how each of these processes can be used to identify a complete set of trials, supporting new methods that use clinical trial registries in systematic reviews.

The objective of this systematic review is to quantify the processes that have been used to link clinical trial registrations to their published results in studies that examined the completeness and accuracy of clinical trial reporting.

METHODS

Inclusion and exclusion criteria

We will include all English-language studies that use one or more of the clinical trial registries included in the WHO International Clinical Trial Platform (WHO ICTRP),³⁸ to compare what was registered with what was published, determine the proportion of published trial reports that have been registered or the proportion of registered trials that have been published. Studies will be excluded if they do not report the number of clinical trials for which they identified links or if the study is describing a trial or reviewing clinical evidence.

Search strategy

Relevant articles will be identified by searching PubMed and EMBASE for studies that meet the inclusion criteria. These databases were selected because they are known to have good coverage of clinical research,^{39 40} and other databases typically used in systematic reviews of clinical evidence (such as the Cochrane Central Register of Controlled Trials) were irrelevant to the topic of the review. Owing to the atypical nature of the included studies, we designed the search strategy with the support of a medical research librarian, and it was designed to balance the number of articles returned by the broad terms covering the clinical trial registries by constraining the search using terms that were common to the set of relevant studies. We considered a set of 50 articles we knew met the inclusion criteria and used their titles, abstracts and keywords to define a search strategy that returned all 50 articles without returning an unmanageable number of irrelevant articles (table 1). We will additionally hand-search the reference lists of all included studies to identify any other articles that may have been missed by our searches. The complete steps for searching both databases are included in the online supplementary material.

Two reviewers will evaluate the articles returned by the searches against the inclusion and exclusion criteria. Duplicate studies will be removed by automatically comparing digital object identifiers (DOIs) across the search results where possible, and by manually evaluating titles, authors and journal names for the remainder. In each of the two phases of screening for eligibility (title/abstract and then full text review), disagreements about inclusion will be resolved by a third author and by discussion, as needed.

Data extraction process

Data from studies will be extracted independently by both reviewers and then compared, reporting the level of agreement for each information element. The

Table 1 Search strategy used for PubMed and EMBASE bibliographic databases

Search terms	
#1	Any registry from WHO ICTRP OR "trial registry"[Title/Abstract] OR "trial register"[Title/Abstract] OR "trial registries"[Title/Abstract] OR "trials registry" [Title/Abstract] OR "registry of clinical trials" [Title/Abstract]
#2	("trial registration"[Title/Abstract] AND (discrepancy[Title/Abstract] OR discrepancies[Title/Abstract] OR consistency[Title/Abstract] OR inconsistency[Title/Abstract]))
#3	#1 OR #2
#4	(unregistered[Title/Abstract] OR non-publication[Title/Abstract] OR nonpublication[Title/Abstract] OR unpublished[Title/Abstract] OR published[Title/Abstract] OR (registered[Title/Abstract] AND (publication[Title/Abstract] OR clinical trial as topic [MeSH Terms]))) Note: For EMBASE, equivalent Emtree is "clinical trial (topic)".
#5	#3 AND #4
#6	("outcome reporting bias"[Title/Abstract] OR "selective reporting"[Title/Abstract] OR "selective outcome reporting"[Title/Abstract] OR "missing outcome data"[Title/Abstract] OR "publication bias"[MeSH Terms] OR ("reporting quality"[Title/Abstract] AND publications[Title/Abstract])) Note: For EMBASE, the publication bias MeSH Term has no equivalent and is removed.
#7	#5 OR #6

ICTRP, International Clinical Trial Registry Platform; MeSH, Medical Subject Headings.

information to be extracted includes the following: (1) the number of trial registry entries examined or identified in the study; (2) the number of published trial reports examined or identified in the study; (3) the trial registry or registries used; (4) the purpose of the study (such as measuring publication bias, outcome reporting bias or the number of published trials that were registered); (5) the application domain; (6) the processes used to identify the links and (7) the proportions of the links found for each method if available.

In relation to the method for identifying the links, we categorise links as one of three types—automatic, inferred or inquired. Automatic links are those for which the unique identifier from the trial registry entry is used to identify links to published results without requiring any inference or manual work. Inferred links are those for which the investigators used the characteristics of the trials to search and reconcile links to or from published trial reports. Inquired links are those confirmed by contacting trial investigators or authors to identify the location of published results.

Data synthesis

Using the information extracted from the articles, we will pool the overall proportions of trials for which links were identified. As they represent different types of links between trial registries and published reports, studies that start from a cohort of trial registry entries and identify published results will be pooled separately from studies that start from a cohort of published trial reports and identify trial registry entries.

Heterogeneity in the overall number and proportion of links identified in these studies is expected to come partially from differences in the processes being used to identify links, the period in which the trials were completed and published (reflecting temporal changes in the policy and practice of trial registration and reporting) and the specific application domains (some conditions or interventions may be more likely to have registered trials published or published trials registered). To account for these differences in the overall pooled estimates, we will estimate the contributions of each of the three categories of linking processes to the overall estimates wherever the information is available. The result will include estimates of the proportions of links that can be automatically captured, the larger proportion that can be reliably identified when investigators search for and infer links, and the larger proportion that can be identified when investigators contact trial investigators for more information.

We are also interested in examining whether the processes for identifying links between registry entries and trial results have changed over time. To measure the differences over time, we will examine the trend in the proportion of links identified overall—as well as using each of the three categories of linking processes—by applying linear regression relative to the midpoints of the data collection periods specified in each of the studies.

DISCUSSION AND DISSEMINATION

To the best of our knowledge, this is first systematic review surveying the processes used to identify links between clinical trial registries and published clinical trial results. By aggregating the results of many trials in the area to estimate the proportions of links that can be identified through automatic linking, inference and inquiry, this systematic review is expected to advance the field in several ways. First, the systematic review will be used to determine whether differences in the processes for identifying links between registry entries and trial reports can partially explain differences in the results of existing studies of publication bias and outcome reporting bias, which in turn may be used to help standardise the way these studies are undertaken in the future. Second, by determining the proportions of trial registry entries that can be automatically linked to their results, we can help guide new systematic review technologies that rely on links to improve methods used in the identification of trials.

One limitation of the systematic review process is the exclusion of studies that are not published in English, which may mean that we miss some articles describing registries based in non-English speaking countries. As a substantial proportion of the studies that will be included are designed for purposes other than simply identifying the links between registries and published results, the description of the processes used to link the two may be limited, and this may limit our ability to determine the proportions of links captured automatically, by inference or by inquiry.

Our findings will be reported on the basis of guidelines from Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement. The results of this review will be submitted for publication in a peer-reviewed medical journal. Other forms of dissemination will include direct engagement with clinical trial registry developers.

Twitter Follow Adam Dunn at @adamgdunn

Contributors RB and AGD developed the systematic review protocol, drafted the manuscript, critically revised the manuscript and approve of the final version.

Funding RB is supported by a Macquarie University Postgraduate Scholarship.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

REFERENCES

1. Dickersin K, Rennie D. Registering Clinical Trials. *JAMA* 2003;290:516–23.
2. McCray AT. Better access to information about clinical trials. *Ann Intern Med* 2000;133:609–14.

3. Manheimer E, Anderson D. Survey of public information about ongoing clinical trials funded by industry: evaluation of completeness and accessibility. *BMJ* 2002;325:528–31.
4. Zarin DA, Tse T, Williams RJ, *et al.* The ClinicalTrials.gov results database—update and key issues. *N Engl J Med* 2011;364:852–60.
5. De Angelis CD, Drazen JM, Frizelle FA, *et al.*, International Committee of Medical Journal Editors. Is this clinical trial fully registered?—A statement from the International Committee of Medical Journal Editors. *N Engl J Med* 2005;352:2436–8.
6. Jones CW, Handler L, Crowell KE, *et al.* Non-publication of large randomized clinical trials: cross sectional analysis. *BMJ* 2013;347:f6104.
7. Riveros C, Dechartres A, Perrodeau E, *et al.* Timing and completeness of trial results posted at ClinicalTrials.gov and published in journals. *PLoS Med* 2013;10:e1001566; discussion e1001566.
8. Dwan K, Gamble C, Williamson PR, *et al.* Systematic review of the empirical evidence of study publication bias and outcome reporting bias—an updated review. *PLoS ONE* 2013;8:e66844.
9. Bourgeois FT, Murthy S, Mandl KD. Outcome reporting among drug trials registered in ClinicalTrials.gov. *Ann Intern Med* 2010;153:158–66.
10. Vawdrey DK, Hripcsak G. Publication bias in clinical trials of electronic health records. *J Biomed Inform* 2013;46:139–41.
11. Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. *J Clin Epidemiol* 2000;53:207–16.
12. Ioannidis JP. Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *JAMA* 1998;279:281–6.
13. Ross JS, Mocanu M, Lampropoulos JF, *et al.* Time to publication among completed clinical trials. *JAMA Intern Med* 2013;173:828–8.
14. Ross JS, Mulvey GK, Hines EM, *et al.* Trial publication after registration in ClinicalTrials.gov: a cross-sectional analysis. *PLoS Med* 2009;6:e1000144.
15. Ross JS, Tse T, Zarin DA, *et al.* Publication of NIH funded trials registered in ClinicalTrials.gov: cross sectional analysis. *BMJ* 2012;344:d7292.
16. Gordon D, Taddei-Peters W, Mascette A, *et al.* Publication of trials funded by The National Heart, Lung, and Blood Institute. *N Engl J Med* 2013;369:1926–34.
17. De Jong JP, Ter Riet G, Willems DL. Two prognostic indicators of the publication rate of clinical studies were available during ethical review. *J Clin Epidemiol* 2010;63:1342–50.
18. Tang E, Ravaut P, Riveros C, *et al.* Comparison of serious adverse events posted at ClinicalTrials.gov and published in corresponding journal articles. *BMC Med* 2015;13:189.
19. Pranić S, Marušić A. Changes to registration elements and results in a cohort of ClinicalTrials.gov trials were not reflected in published articles. *J Clin Epidemiol* 2016;70:26–37.
20. Jones CW, Keil LG, Holland WC, *et al.* Comparison of registered and published outcomes in randomized controlled trials: a systematic review. *BMC Med* 2015;13:282.
21. Dwan K, Altman DG, Cresswell L, *et al.* Comparison of protocols and registry entries to published reports for randomised controlled trials. *Cochrane Database Syst Rev* 2011;(1):MR000031.
22. Kirkham JJ, Dwan KM, Altman DG, *et al.* The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ* 2010;340:c365.
23. Page MJ, McKenzie JE, Kirkham J, *et al.* Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions. *Cochrane Database Syst Rev* 2014;(10):Mr000035.
24. Ewart R, Lausen H, Millian N. Undisclosed changes in outcomes in randomized controlled trials: an observational study. *Ann Fam Med* 2009;7:542–6.
25. De Oliveira GS Jr, Jung MJ, McCarthy RJ. Discrepancies between randomized controlled trial registry entries and content of corresponding manuscripts reported in anesthesiology journals. *Anesth Analg* 2015;121:1030–3.
26. Mathieu S, Boutron I, Moher D, *et al.* Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA* 2009;302:977–84.
27. Walker KF, Stevenson G, Thornton JG. Discrepancies between registration and publication of randomised controlled trials: an observational study. *JRSM Open* 2014;5:2042533313517688.
28. Huser V, Cimino JJ. Linking ClinicalTrials.gov and PubMed to track results of interventional human clinical trials. *PLoS ONE* 2013;8:e68409.
29. Huser V, Cimino JJ. Precision and negative predictive value of links between ClinicalTrials.gov and PubMed. *AMIA Annu Symp Proc* 2012;2012:400–8.
30. Jones CW, Keil LG, Weaver MA, *et al.* Clinical trials registries are under-utilized in the conduct of systematic reviews: a cross-sectional analysis. *Syst Rev* 2014;3:126–6.
31. Wolfe N, Götzsche PC, Bero L. Strategies for obtaining unpublished drug trial data: a qualitative interview study. *Syst Rev* 2013;2:1–11.
32. Schroll JB, Bero L, Götzsche PC. Searching for unpublished data for Cochrane reviews: cross-sectional study. *BMJ* 2013;346:f2231.
33. Garner P, Hopewell S, Chandler J, *et al.*, Panel for updating guidance for systematic reviews (PUGs). When and how to update systematic reviews: consensus and checklist. *BMJ* 2016;354:i3507.
34. Takwoingi Y, Hopewell S, Tovey D, *et al.* A multicomponent decision tool for prioritising the updating of systematic reviews. *BMJ* 2013;347:f7191.
35. Garrity C, Tsertsvadze A, Tricco AC, *et al.* Updating systematic reviews: an international survey. *PLoS ONE* 2010;5:e9914.
36. Cohen AM, Ambert K, McDonagh M. Studying the potential impact of automated document classification on scheduling a systematic review update. *BMC Med Inform Decis Mak* 2012;12:33.
37. O'Mara-Eves A, Thomas J, McNaught J, *et al.* Using text mining for study identification in systematic reviews: a systematic review of current approaches. *Syst Rev* 2015;4:5.
38. Primary Registries [June 2]. <http://www.who.int/ictrp/network/primary/en/>
39. Bramer WM, Giustini D, Kramer BM. Comparing the coverage, recall, and precision of searches for 120 systematic reviews in Embase, MEDLINE, and Google Scholar: a prospective study. *Syst Rev* 2016;5:39.
40. Preston L, Carroll C, Gardois P, *et al.* Improving search efficiency for systematic reviews of diagnostic test accuracy: an exploratory study to assess the viability of limiting to MEDLINE, EMBASE and reference checking. *Syst Rev* 2015;4:82.