



ORIGINAL ARTICLE

Cochrane Centralised Search Service showed high sensitivity identifying randomized controlled trials: A retrospective analysis

A.H. Noel-Storr^{a,*}, G. Dooley^b, S. Wisniewski^{a,f}, J. Glanville^c, J. Thomas^d, S. Cox^e,
R. Featherstone^f, R. Foxlee^f

^aRadcliffe Department of Medicine, Cochrane Dementia and Cognitive Improvement Group, Oxford University, Oxford, UK

^bMetaxis Ltd, Oxford, UK

^cYork Health Economics Consortium, University of York, Oxford, UK

^dEPPI-Centre, University College London, Oxford, UK

^eCochrane ENT, NDS, Oxford University, Oxford, UK

^fCochrane Editorial and Methods Department, Cochrane, London, UK

Accepted 11 August 2020; Published online xxxx

Abstract

Background and objectives: The Cochrane Central Register of Controlled Trials (CENTRAL) is compiled from a number of sources, including PubMed and Embase. Since 2017, we have increased the number of sources feeding into CENTRAL and improved the efficiency of our processes through the use of application programming interfaces, machine learning, and crowdsourcing. Our objectives were twofold: (1) Assess the effectiveness of Cochrane's centralized search and screening processes to correctly identify references to published reports which are eligible for inclusion in Cochrane systematic reviews of randomized controlled trials (RCTs). (2) Identify opportunities to improve the performance of Cochrane's centralized search and screening processes to identify references to eligible trials.

Methods: We identified all references to RCTs (either published journal articles or trial registration records) with a publication or registration date between 1st January 2017 and 31st December 2018 that had been included in a Cochrane intervention review. We then viewed an audit trail for each included reference to determine if it had been identified by our centralized search process and subsequently added to CENTRAL.

Results: We identified 650 references to included studies with a publication year of 2017 or 2018. Of those, 634 (97.5%) had been captured by Cochrane's Centralised Search Service. Sixteen references had been missed by the Cochrane's Centralised Search Service: six had PubMed-not-MEDLINE status, four were missed by the centralized Embase search, three had been misclassified by Cochrane Crowd, one was from a journal not indexed in MEDLINE or Embase, one had only been added to Embase in 2019, and one reference had been rejected by the automated RCT machine learning classifier. Of the sixteen missed references, eight were the main or only publication to the trial in the review in which it had been included.

Conclusion: This analysis has shown that Cochrane's centralized search and screening processes are highly sensitive. It has also helped us to understand better why some references to eligible RCTs have been missed. The CSS is playing a critical role in helping to populate CENTRAL and is moving us toward making CENTRAL a comprehensive repository of RCTs. © 2020 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Declaration of interest: All authors are on the Centralised Search Service project team.

Funding: The Centralised Search Service project is funded by the Editorial and Methods Department in Cochrane (London, UK). This analysis was carried out as part of that project.

Declarations of interest: The authors are all members of the Centralised Search Service project team. The Centralised Search Service is an initiative funded by the Cochrane Editorial and Methods Department. Anna Noel-Storr is an information specialist with the Cochrane Dementia Group but is seconded to Cochrane HQ 3 days per week. Under this secondment, Anna leads the Cochrane Crowd initiative and works on the Centralised Search Service initiative. Gordon Dooley is Managing Director at Metaxis Ltd and is contracted by Cochrane to develop the Cochrane Register of Studies and Cochrane Crowd. Susanna Wisniewski is employed as a project officer for the Centralised Search Service initiative. Julie Glanville was contracted by Cochrane to develop methodological filters used by the Centralised Search Service initiative. James Thomas developed the RCT classifiers used by the Centralised Search Service initiative. Samantha Cox is an information specialist for Cochrane ENT Group and is seconded 1 day per week to the Cochrane Information Specialist Support Team. Robin Featherstone is an information specialist for the Cochrane Editorial Methods Department. Ruth Foxlee is the Senior Programme Manager for the Cochrane Editorial Methods Department and has oversight of the CSS initiative.

* Corresponding author. Anna Noel-Storr, Radcliffe Department of Medicine, John Radcliffe Hospital, Oxford University, Room 4401c, Level 4, Oxford, UK, OX3 9DU.

E-mail address: anna.noel-storr@rdm.ox.ac.uk (A.H. Noel-Storr).

Keywords: Cochrane central register of controlled trials; Randomized controlled trial; Information retrieval; Methodological filter; Crowdsourcing; Machine learning; Evidence synthesis; Systematic review

1. Introduction

The Cochrane Central Register of Controlled Trials (CENTRAL) is a bibliographic database populated with reports of randomized and quasi-randomized controlled trials (RCTs and q-RCTs) [1,2]. CENTRAL is available through the Cochrane Library. Most review teams, whether they are producing Cochrane or non-Cochrane reviews, can access CENTRAL for free, either through national licenses or institutional subscriptions. Reports of RCTs are added to CENTRAL through two main routes: (1) via Cochrane Information Specialists identifying and manually adding trial records and (2) by a centralized search initiative, called the Centralised Search Service (CSS), managed by Cochrane's Editorial and Methods Department.

At the time of writing (January 2019), there are five sources that are searched centrally: PubMed and [ClinicalTrials.gov](https://clinicaltrials.gov), both produced by the US National Library of Medicine; [Embase.com](https://embase.com) produced by Elsevier; the World Health Organization's International Clinical Trials Registry Platform (ICTRP); and KoreaMed produced by the Korean Association of Medical Journal Editors. The service is also adding a sixth source: The Cumulative Index to Nursing and Allied Health Literature hosted by EBSCOhost. The Cumulative Index to Nursing and Allied Health Literature records are expected to appear in CENTRAL at the end of the first quarter of 2020. For each source, we have developed bespoke workflows with the aim of capturing all possible reports of RCTs and q-RCTs.

The CSS uses four main approaches to identify relevant records. Not all are used for each of the sources covered (a summary of the overall workflow is given in [Table 1](#) and depicted in [Fig. 1](#)). The four approaches are:

1. Direct feed
2. Sensitive search
3. Machine learning
4. Crowdsourcing

1.1. Direct feed

The first approach is the “direct feed” which consists of records that have been indexed in the source databases as RCTs. Wherever possible, we aim to identify potential “direct feeds” of records reporting an RCT into CENTRAL. This route is the most efficient approach because it does not require any manual assessment/screening of records. We currently have “direct feeds” in place for four of the five sources: PubMed, Embase, [ClinicalTrials.gov](https://clinicaltrials.gov), and ICTRP. However, the “direct feeds” only capture a proportion of the eligible records from each source. Other approaches are therefore needed to identify the remaining RCTs.

1.2. Sensitive search

The second approach is to use a sensitive search. Records from all sources which cannot be identified through the “direct feed” (i.e., they do not have the required index terms) are identified through a search which has been developed for each source [3,4]. As the results from a sensitive search for RCTs inevitably contain many nonrelevant records, additional checks are then required to ensure that only randomized study reports are retained. These additional checks are in two phases: first, records are passed through a machine learning classifier to eliminate clearly irrelevant records [5]; and second, the remainder, are checked by Cochrane Crowd [6].

1.3. Machine learning

The third approach, which supplements the sensitive search, uses machine learning. The automated machine learning classifiers provide likelihood scores as to whether the record is describing an RCT. The CSS uses two machine learning classifiers, one developed for the bibliographic records such as those identified from Embase, and one developed for trial registry records from [ClinicalTrials.gov](https://clinicaltrials.gov). For more detail on the training, calibration, and validation of the bibliographic RCT classifier, see the study by Thomas [5]. The RCT machine learning classifiers are currently used to remove “noise” (nonrelevant records) from large record sets. In other words, we are not using the RCT classifiers to identify RCTs with high precision; we are using them to remove the very obvious non-RCT records, thereby reducing the amount of manual screening required.

1.4. Crowdsourcing

The final component in the workflow is “the Crowd”. Records that have not been accounted for in either the direct feeds or excluded by the RCT classifiers need to be manually screened. These records are sent to Cochrane Crowd (<https://crowd.cochrane.org>), Cochrane's citizen science platform that hosts tasks aimed at identifying particular types of health research. Cochrane Crowd is open to all but contributors must first complete a brief training module before being able to screen live records. In addition to training, the Crowd approach uses an agreement algorithm to help ensure collective accuracy of the data output. The current algorithm in place for the identification of RCTs from bibliographic sources requires that each record needs four consecutive agreement classifications for that record to be deemed either an RCT or not an RCT. The current agreement algorithm in place for the identification of

What is new?**Key findings**

- Cochrane's Centralised Search Service is performing well with a recall of 97.5%.

What this adds to what was known?

- The service uses machine learning and crowdsourcing for the identification of randomised trials.

What is the implication and what should change now?

- Restricting searches to CENTRAL might be adequate in certain contexts.

RCTs from trial registries (i.e., the records from ClinicalTrials.gov and ICTRP) is that each requires three consecutive agreement classifications for that record to be deemed either an RCT or not an RCT. Disagreeing classifications or records that receive an unsure classification go to “resolver” screeners in the Cochrane Crowd. Resolvers are highly experienced screeners who are tasked with making a final decision on records that have not received the required consecutive agreement classifications. For more detailed

information regarding the agreement algorithms used and the accuracy of this crowdsourced approach, see the study by Noel-Storr [6]. The Cochrane Crowd community stands at over 17,000 people from over 150 countries.

2. Aims and objectives

Our aim was to evaluate Cochrane's CSSs to assess its effectiveness at capturing reports of randomized trials for Cochrane intervention reviews. We sought to determine overall comprehensiveness and to assess the performance of each component of the workflow. We were concerned specifically with sensitivity: that is, establishing whether our processes identify all the studies that they were designed to, rather than evaluating their efficiency in terms of their specificity.

We also sought to analyze in detail the reasons why references to studies were not captured by the CSS and to recommend any improvements to our workflows and processes that we could identify.

3. Methods

We conducted a retrospective analysis of Cochrane intervention reviews available in March 2019 and downloaded references to their included studies that had a publication

Table 1. Individual workflows for centrally searched sources as of December 2019. More detailed information on the current workflows in place can be found at <http://www.cochranelibrary.com/help/central-creation-details.html>

Source (provider)	Workflow description	Harvested from external source
PubMed (National Library of Medicine)	Direct feed of records into CENTRAL based on index terms: <i>"randomized controlled trial"[Publication Type] OR "controlled clinical trial"[Publication Type]</i>	Monthly API call on 16th of each month
Embase (Elsevier)	Direct feed of records into CENTRAL based on Emtree term: <i>Randomized controlled trial</i> Sensitive search of Embase.com via the Embase.com API; the results sent to RCT classifier, and the remaining records sent to Cochrane Crowd for manual screening	Monthly API call on 15th of each month
ClinicalTrials.gov (National Library of Medicine)	Direct feed of records into CENTRAL of all records with randomized controlled trial in study design field Download all other records; the results sent to classifier. Those meeting threshold criteria are then sent to Cochrane Crowd for manual screening	Daily API call
ICTRP (World Health Organisation)	Download all records; remove CT.gov records. The remaining records sent to Cochrane Crowd for manual screening.	Monthly API call on 15th of each month
KoreaMed (Korean Association of Medical Journal Editors)	Download all records; records are sent to Cochrane Crowd for manual screening.	Monthly API call on 15th of each month
CINAHL (EBSCOhost)	Sensitive search of CINAHL via API; the results sent to RCT classifier, and the remaining records sent to Cochrane Crowd for manual screening	Daily API call from November 2020

By combining these approaches—API direct feeds, sensitive searches, machine learning classifiers, and crowdsourcing manual screening via Cochrane Crowd—the CSS has established an effective process for identifying RCTs for CENTRAL. This RCT identification workflow required evaluation to ensure and improve efficiency and accuracy.

API: application programming interfaces; CINAHL, The Cumulative Index to Nursing and Allied Health Literature.

(or trial registration) date of either 2017 or 2018. We chose these 2 years because they are the two most recent years where we have had the CSS operating. We are able to use this data set because at present, in the vast majority of cases, studies included in Cochrane reviews are not identified from a single search of CENTRAL but through extensive and sensitive searches conducted across multiple sources in accordance with Methodological Expectations for Cochrane Intervention Reviews [7] and the Cochrane Handbook [8]. If studies had been identified through searches of CENTRAL only, we would not have been able to ascertain the comprehensiveness of CSS processes.

After downloading all the 2017 and 2018 included references to studies, we removed duplicate references and

references to nonrandomized studies. We identified these studies by examining the inclusion criteria for each review. If the review stated that it had included study designs other than randomized controlled trials, we then checked the characteristics of included studies table within the review to discern whether the included studies were RCTs or not. Two assessors working independently then categorized each reference as per the following: 1) *journal article* (including letters, errata etc.), 2) *conference publication*, 3) *trial registry record*, and 4) *other* for record types not covered by the CSS, for example, clinical study reports and email correspondence.

We also noted whether the reference had been flagged as the primary reference to an RCT or a secondary publication by the individual review author teams because trialists very

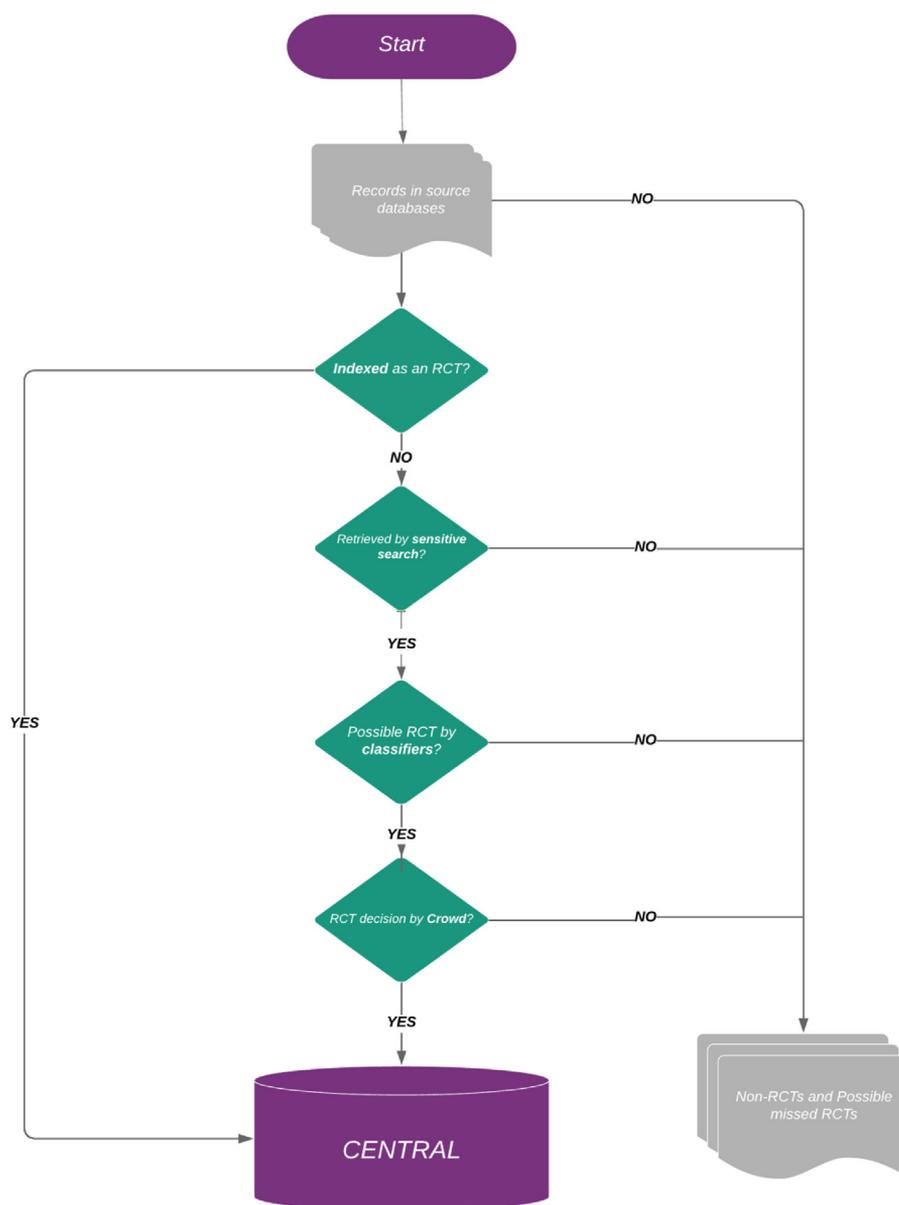


Fig. 1. Study identification workflow.

often produce more than one publication or research output for a single trial [9].

With this categorized data set, we then constructed an audit trail for each record to ascertain whether it had been identified by the CSS and, if so, through which approach. For example, whether the reference had been picked up by the CSS via a direct feed or via a sensitive search and crowdsourcing. We did not assess whether the references to included studies were retrieved by the actual searches performed in CENTRAL for the reviews. While this is an important question, it goes beyond the scope of this evaluation which sought to assess recall in terms of whether the centralized processes identified the RCTs included in reviews. We used a relative recall approach often used in studies evaluating the performance of methodological search filters [10]. This approach uses a set of known relevant records (the included studies) as its denominator, rather than a handsearched gold standard data set.

4. Results

We retrieved 782 references to included studies from 274 reviews with a publication year of 2017 or 2018. After

removing the duplicates and the non-RCT records, we were left with 739 records. Figure 2 shows the flow of references used in this analysis and the breakdown of record type based on the categories we used.

The 650 references to included studies from 262 Cochrane reviews were record types covered by the CSS. We reviewed the methods section of a random sample of 25% (65) of the 262 reviews to check that multiple sources (i.e., not just CENTRAL) had been searched. Within this sample, two search approaches were described: 1. searches developed and carried out across multiple sources specifically for the review and 2. searches carried out in specialized registers of the review group responsible for the review. For those reviews reliant on register-only searches, we checked that the searches run for the maintenance of the register were across multiple sources. 51 of the 65 reviews checked reported carrying out bespoke searches across multiple sources specifically for the review; 14 reported using their specialized register as the main source searched. None of the review methods checked reported searching only CENTRAL.

Of the 650 references to RCTs included in Cochrane reviews, 97.5% (634) had been captured by the centralized study identification processes. The majority of these had

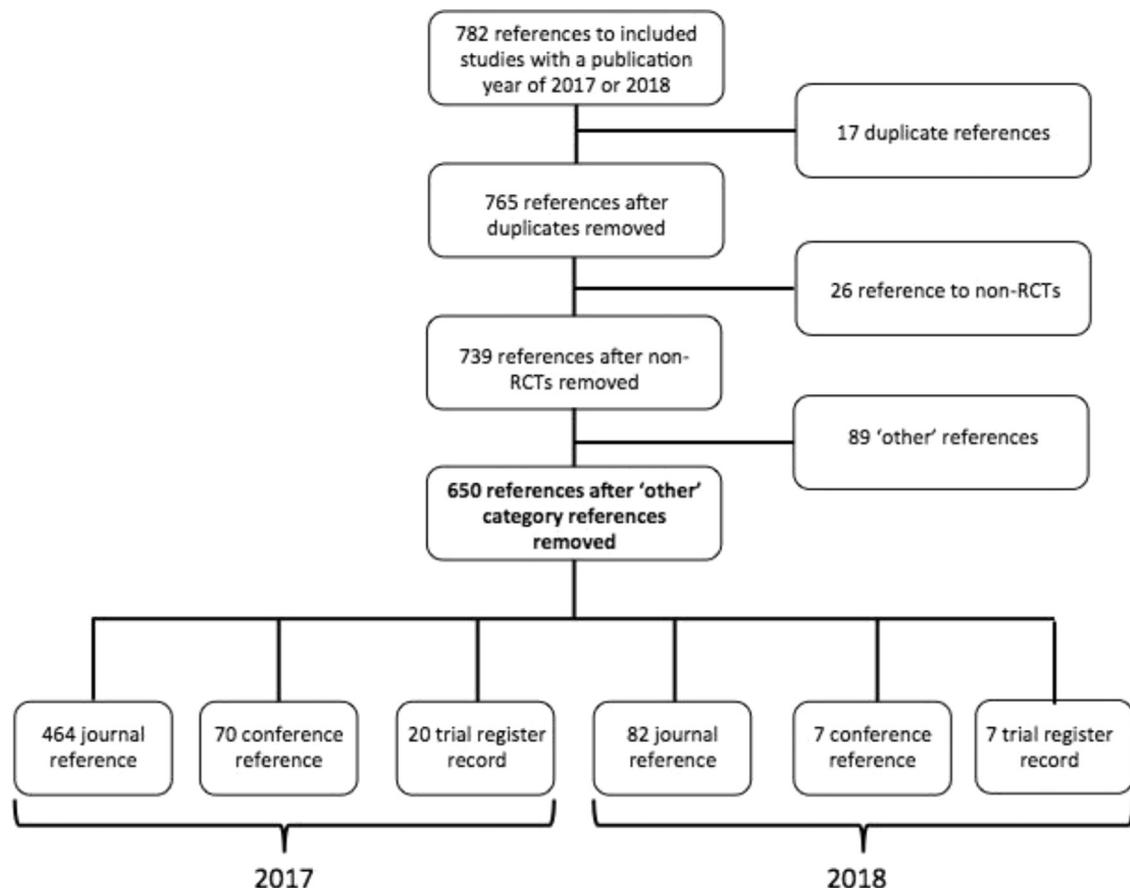


Fig. 2. Flow diagram of references to studies included in this retrospective analysis.

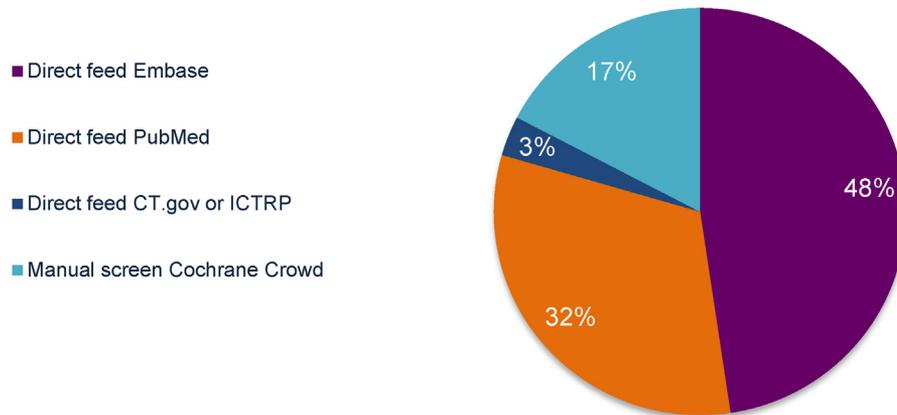


Fig. 3. Breakdown of RCTs identified by the CSS approach.

been identified by the PubMed and Embase direct feeds of records 32% (202) and 48% (302), respectively. A further 110 (17%) references to included studies had been identified by Cochrane Crowd, and 20 (3%) had come in through the direct feeds of trial registry records (see Fig. 3).

Sixteen references to included studies were not identified through the CSS. Of these, six (38%) were references in PubMed, but they did not have the RCT or controlled clinical trial publication type index term and so were not picked up by the PubMed direct feed or by the Embase direct feed or sensitive search. While the vast majority of PubMed records are in MEDLINE and therefore identifiable from Embase (which subsumed MEDLINE content in 2011), some records with PubMed-not-MEDLINE status remain outside of the main data set. The two main reasons for records acquiring a PubMed-not-MEDLINE status are (1) they have yet to be indexed for MEDLINE or (2) they are records to articles in journals not covered by MEDLINE or are in the National Library of Medicine's PubMed Central open archive of full text journal articles. Some of these PubMed-not-MEDLINE status records will become retrievable from Embase over time; however, others may not. All six missed publications were the primary, and only, study records listed in the reviews for those trials.

Of the remaining ten missed references, three had been identified by the sensitive Embase search but had then been incorrectly rejected as non-RCTs by the Cochrane Crowd. The three references were as follows: a long-term follow-up report of an RCT, a letter, published in a journal, about an RCT, and an analysis of a secondary outcome of an RCT. All three were secondary publications to the trial they were describing.

A further four references had been missed by the sensitive Embase search. Of these, one was the only reference for that included study; the other three were secondary publications (i.e., there were other references to those trials included in the review). For each of the missed references, the titles, abstracts, and index terms contained no explicit description to indicate they were reporting or describing

an RCT. One was a subgroup analysis to an RCT where the name of a trial was provided, but there were no other descriptors that indicated that the trial was an RCT. Another was a letter published in a journal about a trial. This Embase record contained only the title and none of the index terms related to study design. The third missed secondary publication was a long-term follow-up of an RCT. The final missed reference described a controlled study but did not provide details on how the participants had been allocated to each arm of the trial. The abstract described the trial's aim as examining the "comparative efficacy" of a 12 week treatment program versus a "treatment as usual" group and was indexed with the Emtree headings *controlled study* and *comparative effectiveness*. We currently use the narrower Emtree term *controlled clinical study* in the Cochrane sensitive Embase search, rather than *controlled study*; therefore, despite the sensitivity of the Cochrane Embase—centralized search, this reference was not captured.

The machine learning classifier was identified as the cause of one missing reference. The RCT classifier works to remove the records with a very low probability of describing an RCT. In other words, it handles many of the clear-cut non-RCTs, thereby freeing up human effort (the Crowd) to manually screen the records that would challenge the machine classifier. The expected recall rate of the classifier is around 99.5% on studies included in Cochrane reviews [5]. Therefore only missing 1 study is exceeding this expected performance. The missed study was a secondary publication of a RCT that assessed biomarker data available from a subset of the original trial's participants. With the exception of the word "randomized" being used once in the abstract, there were no other indications that this report was related to a randomized trial.

The final two missed references were a conference publication to a 2017 study that was not added to Embase until week 34 of 2019 and not retrieved by the feeds during the period of interest and a reference in a journal not indexed by any of the sources covered by the CSS: *Modern*

Approaches in Drug Designing. The former missed reference was a secondary one, and the latter was flagged as the primary reference to the included study.

5. Discussion

This analysis found that the CSS—a Cochrane initiative that aims to identify as many reports of RCTs as efficiently and as accurately as possible through a combination of direct feeds, sensitive searching, crowdsourcing, and machine learning—is achieving high sensitivity. While some studies were missed through these approaches, the number missed was small and comparable with the expected recall of traditional methodological filters and the screening of abstracts by review author teams [11]. Only 2.5% of references to included studies in our test set were not picked up by the CSS, and of those 16 missed references, only eight (50%) were flagged as the primary article to the RCT in the Cochrane review. In addition, this analysis has shown the valuable role of Cochrane Crowd, which identified 17% of the references.

In terms of the range of sources we currently search as part of the CSS initiative, this analysis also indicates that Cochrane's coverage of English-language journal articles, conference publications, and trial register records is highly sensitive. Importantly, only one reference to a journal article included in a Cochrane review was missed because it was not indexed in any of the bibliographic databases covered by the CSS. This is helpful information in terms of prioritizing which sources should be the next focus for any centralized searching efforts. However, the fact there was only one missed study could indicate that searches for Cochrane reviews are potentially not broad enough in terms of less mainstream databases and non-English language sources. Options for future objectives of the CSS could be to target non-English language material and other record types such as clinical study reports [12,13].

This analysis also provided us with a better understanding of what we can do to improve our current processes and some of that work has already begun. For example, we have revised the Cochrane Crowd Quick Reference Guide to make clear that follow-up studies to RCTs are to be selected for CENTRAL. We also now require that Crowd contributors repeat the training module every 6 months to remind them of the inclusion criteria for CENTRAL. We are reviewing the Embase sensitive search to see whether it should be amended slightly in light of the few missed references, and we are currently evaluating the existing PubMed RCT filters to capture those references that are in PubMed but have not been indexed with the RCT publication type term. We have also recently updated the RCT classifier and tested whether references rejected by the old version would now be included.

One question frequently posed to the CSS team is whether searches for randomized evidence can now be

limited to searching only CENTRAL. Several research articles have sought to evaluate the comprehensiveness of other major bibliographic databases [14–17] or trial registries [18]. This analysis indicates that the vast majority of published articles, conference proceedings, and trial registry records are being successfully identified by the centralized searching and screening processes. However, there are a number of factors that should be taken into consideration when deciding which sources to search and, specifically, whether there is still a need to search the source databases currently covered by the Cochrane CSS. We will start first with specific limitations of this analysis before describing a number of more general factors that could help inform decisions about which sources to search.

5.1. Limitations of this analysis

This analysis has focused on a very specific time frame: studies with a publication year of 2017 or 2018, therefore our results are limited to more recent reports of randomized trials. The reporting of randomized trials has likely improved over time due to the CONSORT initiative [19,20]; this may have made identifying randomized trials easier. Most new reviews would normally plan to search for trials over all years and not just those published more recently. This analysis does not help to answer the question of whether someone looking for trials across all dates by just searching CENTRAL would be likely to find them all (or even 97.5% of them). Another limitation is that our sample size for trial registry records is small; therefore the findings of this study should be viewed with caution in relation to this record type.

There are other broader factors to take into consideration when deciding which sources to search, specifically with regard to limiting a search to CENTRAL.

5.2. Time-lag from source database to CENTRAL

Currently, the shortest time possible for a record in a source database to appear in CENTRAL after identification from a source database is between three to 4 weeks. This is because the source databases are currently queried once every month, and CENTRAL is updated once every month. However, some records can take much longer to reach CENTRAL. These are records that need to be resolved in Cochrane Crowd (an average of 11.3% of records across the three Crowd tasks that feed CENTRAL need resolving), either because the Crowd has disagreed in their classifications of a record or has classified a record with an unsure classification. These records can take time to receive a final classification. Resolver screeners, members of the Cochrane Crowd community tasked with making a final decision on records that need resolving, are few in number due to the expertise level required and often have to obtain the full text to make a final decision. This issue, however, does raise the question around what would be considered

acceptable levels of ineligible records being submitted to CENTRAL. This analysis has focused entirely on the sensitivity of current processes. However, we do know that some ineligible records reach CENTRAL via the direct feeds and via Cochrane Crowd. A further small drop in specificity may be acceptable if it enabled faster delivery of these RCTs into CENTRAL.

5.3. Different versions of centralized searches and processes

This analysis has focused on records retrieved by the most recent version of the searches and processes in place for the CSS. However, these searches and processes have evolved over time. For example, records identified for CENTRAL from Embase were identified on the basis of a different search strategy pre-2014. The latest search strategy in use by the CSS is considered to be more sensitive than previous iterations. It is therefore feasible that a higher proportion of RCTs may have been missed by older, less sensitive searches. Similarly, when new sources are added to the CSS process, there is often a large initial set of records for all years to process, after which monthly processing is quicker. Strategies to deal with large backlogs often differ slightly from the process used to deal with the prospective data feed for the same source. In taking a pragmatic approach to managing backlogs—which we must do because of resource constraints—it is possible that, despite our best efforts, some eligible records may have been lost.

5.4. The search interface

Another consideration for those interested in restricting their searching to CENTRAL is the difference in the search interface and search capabilities in CENTRAL compared with those of the source databases. Only a subset of meta-data is harvested for CENTRAL, so supplementary searches of source databases may yield additional records. This is particularly relevant to the trial registry records where often much more information is available to search within the regional and international registries [21]. We hope to conduct a further analysis using the CENTRAL search strategies reported in the Cochrane reviews we used for this study and to assess whether the trials were successfully captured by those strategies instead or as well as by the searches run directly in the source databases.

5.5. Inclusion criteria for CENTRAL

Study designs that are eligible for CENTRAL have not changed for many years and remain: RCTs and q-RCTs, controlled before and after studies and interrupted time series. The centralized search processes were designed to capture RCTs and q-RCTs controlled trials. We do not currently have any centralized processes in place to identify controlled before and after studies or interrupted time series. In addition, while criteria in terms of study design have been stable

for some time, over the last few years (since 2014), the types of reports eligible for CENTRAL has broadened. For example, post hoc and secondary analyses of RCTs are now included in CENTRAL. The expanded eligibility criteria have implications particularly for those seeking all publications relevant to a single randomized trial, rather than just the main or primary publication.

5.6. Limitations of search strategies

The effectiveness of database retrieval is also impacted by the quality of the search strategies used to search the database. If the only database to be searched is CENTRAL, then the quality of the single search strategy becomes crucial to the success of the review. A more conservative approach of searching a range of databases with different search translations may increase the chances to retrieve relevant records.

5.7. Searching for what purpose?

The final factor to consider, and perhaps the most obvious one, concerns the objective of the search itself. For example, rapid reviews or scoping searches may accept lower sensitivity in favor of precision, whereas searches for Cochrane intervention reviews and living systematic reviews [22] will be primarily concerned with maximizing sensitivity. So, searchers conducting rapid reviews or scoping reviews may be content to use only CENTRAL with a highly sensitive strategy, whereas searchers populating full systematic reviews may wish to search beyond CENTRAL for the reasons discussed above. Context will therefore always be an important consideration.

There are numerous factors that information specialists should consider when deciding which sources to include in their searches. To help inform this decision-making, we present our methods for identifying trial records for CENTRAL transparently and completely.

6. Conclusion

The CSS has established processes for identifying RCTs for inclusion in CENTRAL by using a combination of API direct feeds, sensitive searches, machine learning classifiers, and crowdsourced manual screening via Cochrane Crowd. Our evaluation has found that the workflow achieves a very high level of sensitivity. We have also identified ways to improve the CSS. We present our process and the results of this evaluation in an effort to support the decision-making of information specialists seeking the best source databases for their work. Although highly sensitive in its coverage, CENTRAL may not yet be seen as a comprehensive source of all relevant trials for systematic reviews for all purposes, and it may, however, be comprehensive enough for some use cases such as searchers undertaking rapid or scoping reviews. In these cases, it will be

important that the quality of the search itself is high and takes into account the limitations discussed. Our processes for identifying RCTs for CENTRAL will continue to evolve through the use of machine learning and the contribution of the Cochrane Crowd community. During these transformations, we will continue to share the results of our process evaluations and our methods for identifying RCTs for CENTRAL.

CRediT authorship contribution statement

A.H. Noel-Storr: Conceptualization, Methodology, Investigation, Data curation. **G. Dooley:** Methodology, Resources, Data curation. **S. Wisniewski:** Methodology, Data curation. **J. Glanville:** Methodology, Visualization, Writing - review & editing. **J. Thomas:** Conceptualization, Methodology, Visualization, Writing - review & editing. **S. Cox:** Methodology, Writing - review & editing. **R. Featherstone:** Methodology, Visualization, Writing - review & editing. **R. Foxlee:** Conceptualization, Methodology, Writing - review & editing.

Acknowledgments

We would like to acknowledge the Cochrane Crowd, an international community of volunteers who help to identify reports of RCTs via the Cochrane Crowd platform.

References

- [1] McKenzie JE, Brennan SE, Ryan RE, Thomson HJ, Johnston RV, Thomas J. Chapter 3: defining the criteria for including studies and how they will be grouped for the synthesis. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, editors. *Cochrane Handbook for Systematic Reviews of Interventions* version 6.0. Available at www.training.cochrane.org/handbook. Accessed September 14, 2020.
- [2] The Cochrane Library. CENTRAL Creation: how CENTRAL is created. Available at <https://www.cochranelibrary.com/central/central-creation>. Accessed September 14, 2020.
- [3] Glanville J, Dooley G, Wisniewski S, Foxlee R, Noel-Storr A. Development of a search filter to identify reports of controlled clinical trials within CINAHL Plus. *Health Info Libr J* 2019;36:73–90. Accessed September 14, 2020.
- [4] Glanville J, Foxlee R, Wisniewski S, Noel-Storr A, Edwards M, Dooley G. Translating the Cochrane EMBASE RCT filter from the Ovid interface to Embase.com: a case study. *Health Info Libr J* 2019;36:264–77.
- [5] Thomas J, McDonald S, Noel-Storr A, Shemilt I, Elliott J, Mavergames C, et al. Machine learning reduces workload with minimal risk of missing studies: development and evaluation of an RCT classifier for Cochrane Reviews. *J Clin Epidemiol*; 2020: Submitted for publication.
- [6] Noel-Storr AH, Dooley G, Elliott J, Steele E, Shemilt I, Mavergames C, et al. An evaluation of Cochrane Crowd finds that crowdsourcing can help to address the challenge of information overload in evidence synthesis. *J Clin Epidemiol*; 2020: Submitted for publication.
- [7] Higgins JPT, Lasserson T, Chandler J, Tovey D, Thomas J, Fleming E, et al. *Methodological Expectations of Cochrane Intervention Reviews*. London: Cochrane; 2019.
- [8] Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. Available at www.training.cochrane.org/handbook. Accessed September 14, 2020.
- [9] Ebrahim S, Montoya L, Kamal el Din M, et al. Randomized trials are frequently fragmented in multiple secondary publications. *J Clin Epidemiol* 2016;79:130–9.
- [10] Sampson M, Zhang L, Morrison A, Barrowman NJ, Clifford TJ, Platt RW, et al. An alternative to the hand searching gold standard: validating methodological search filters using relative recall. *BMC Med Res Methodol* 2006;6.
- [11] McKibbin KA, Wilczynski NL, Haynes RB. Retrieving randomized controlled trials from Medline: a comparison of 38 published search filters. *Health Info Libr J* 2009;26:187–202.
- [12] Doshi P, Jefferson T. Clinical study reports of randomised controlled trials: an exploratory review of previously confidential industry reports. *EMJ Open* 2013;3:e002496.
- [13] Jefferson T, Doshi P, Boutron I, Golder S, Heneghan C, Hodgkinson A, et al. When to include clinical study reports and regulatory documents in systematic reviews. *BMJ Evid Based Med* 2018;23:210–7.
- [14] Halladay CW, Trikalinos TA, Schmid IT, Schmid CH, Dahabreh IJ. Using data sources beyond PubMed has a modest impact on the results of systematic reviews of therapeutic interventions. *J Clin Epidemiol* 2015;68:1076–84.
- [15] Hartling L, Featherstone R, Nuspl M, Shave K, Dryden DM, Vandermeer B. The contribution of databases to the results of systematic reviews: a cross-sectional study. *BMC Med Res Methodol* 2016;16:127.
- [16] Sampson M, Barrowman NJ, Moher D, Klassen TP, Pham B, Platt R, et al. Should meta-analysts search Embase in addition to Medline? *J Clin Epidemiol* 2003;56:943–55.
- [17] Sampson M, de Bruijn B, Urquhart C, Shojania K. Complementary approaches to searching MEDLINE may be sufficient for updating systematic reviews. *J Clin Epidemiol* 2016;78:108–15.
- [18] Baudard M, Yavchitz A, Ravaud P, Perrodeau E, Boutron I. Impact of searching clinical trial registries in systematic reviews of pharmaceutical treatments: methodological systematic review and reanalysis of meta-analyses. *BMJ* 2017;356:j448.
- [19] Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001.
- [20] Hopewell S, Ravaud P, Baron G, Boutron I. Effect of editors' implementation of CONSORT guidelines on the reporting of abstracts in high impact medical journals: interrupted time series analysis. *BMJ* 2012;344:e4178.
- [21] Isojarvi J, Wood H, Lefebvre C, Glanville J. Challenges of identifying unpublished data from clinical trials: getting the best out of clinical trials registers and other novel sources. *Res Synth Methods* 2018;561–78.
- [22] Elliott JH, Synnot A, Turner T, Simmonds M, Akl EA, McDonald S, et al. Living Systematic Review Network. Living systematic review: 1. Introduction—the why, what, when, and how. *J Clin Epidemiol* 2017;91:23–30.