BMJ Open

Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry

Rohit Borah,[1,2] Andrew W Brown,[2,3] Patrice L Capers,[2,3] Kathryn A Kaiser[2,3]

[1]Science and Technology Honors Program, University of Alabama at Birmingham, Birmingham, Alabama, USA
[2]Office of Energetics, Dean's Office, School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama, USA
[3]Nutrition Obesity Research Center, University of Alabama at Birmingham, Birmingham, Alabama, USA

Correspondence to
Dr Kathryn A Kaiser;
kakaiser@uab.edu

ABSTRACT

Objectives: To summarise logistical aspects of recently completed systematic reviews that were registered in the International Prospective Register of Systematic Reviews (PROSPERO) registry to quantify the time and resources required to complete such projects.

Design: Meta-analysis.

Data sources and study selection: All of the 195 registered and completed reviews (status from the PROSPERO registry) with associated publications at the time of our search (1 July 2014).

Data extraction: All authors extracted data using registry entries and publication information related to the data sources used, the number of initially retrieved citations, the final number of included studies, the time between registration date to publication date and number of authors involved for completion of each publication. Information related to funding and geographical location was also recorded when reported.

Results: The mean estimated time to complete the project and publish the review was 67.3 weeks (IQR=42). The number of studies found in the literature searches ranged from 27 to 92 020; the mean yield rate of included studies was 2.94% (IQR=2.5); and the mean number of authors per review was 5, SD=3. Funded reviews took significantly longer to complete and publish (mean=42 vs 26 weeks) and involved more authors and team members (mean=6.8 vs 4.8 people) than those that did not report funding (both p<0.001).

Conclusions: Systematic reviews presently take much time and require large amounts of human resources. In the light of the ever-increasing volume of published studies, application of existing computing and informatics technology should be applied to decrease this time and resource burden. We discuss recently published guidelines that provide a framework to make finding and accessing relevant literature less burdensome.

Strengths and limitations of this study

- This study provides an updated estimate of the time and effort required to conduct and publish a systematic review using a large sample of recently published reviews across a variety of topics of medical interventions.
- The study is limited by incomplete reporting in both sources of the data—the review registry and the published articles.
- These data sources were cross-checked as data were available and conservative assumptions were stated and applied where necessary.

INTRODUCTION

The systematic review can be an effective and scientific method used across disciplines to consolidate vast amounts of research on a specific topic. Over the past 20 years, publishing of systematic reviews has increased exponentially, as has the primary literature.[1] Since the body of primary literature has increased, conducting systematic reviews and meta-analyses has become a necessity to more accurately and comprehensively present accumulated knowledge to scientific, clinical and general audiences. In fact, new mandates are being formulated to require documented systematic reviews as part of research proposals[2] and clinical trial data sharing,[3] both of which have strong implications for how scientific literature and associated data are reported, managed and curated.[4]

The Cochrane Database of Systematic Reviews[5] provides important guidelines and methods for systematic reviews through its *Cochrane Handbook*,[6] which states four key considerations to define prior to beginning: the question, the inclusion and exclusion criteria, the search strategy and the methods. In other words, systematic reviews should review the scientific literature in a scientific manner. The Cochrane Collaboration has built its reputation as the leading resource for systematic reviews by requiring that all Cochrane reviews be "…updated regularly in an effort to ensure that the most recent evidence is incorporated."[7] However, this

onerous requirement may have the side effect of limiting the topics that are created and kept current.

With the ever-growing numbers of journals and publications, an immense amount of time and effort is needed to search the literature and summarise the findings. This is reflected in the Cochrane Collaboration statement that "[h]istorically, the aim was to update Cochrane reviews every two years, but recently there has been a move away from this policy in favour of prioritising the most clinically important reviews for updating."[7] When beginning a new systematic review, authors are faced with the possibility of finding few to no studies that meet their criteria. While finding no studies that meet the criteria can be informative per se, such as by identifying directions for future research, the time and effort required to reach this conclusion may be great. Conversely, the scope of some reviews can be unpredictably large, and it may be difficult to plan the person-hours required to complete the research. The magnitude of this uncertainty has not been defined until now.

Recent efforts, such as the International Prospective Register of Systematic Reviews (PROSPERO)[8] hosted through the University of York's Centre for Reviews and Dissemination, have been deployed to prospectively and centrally register systematic reviews and meta-analyses. The PROSPERO registry "…aims to provide a comprehensive listing of systematic reviews registered at inception to help avoid unplanned duplication and enable comparison of reported review methods with what was planned in the protocol."[8] We found that such a database provides rich data for summarising various logistical aspects of a sample of recently registered and completed systematic reviews. In the present meta-analysis, we aimed to use the PROSPERO registry to quantify the time and person-hours needed to conduct a systematic review. Our specific research questions were as follows:

1. How many people were involved in conducting or authoring the review?
2. How much time was required to complete and publish the review?
3. What was the average efficiency of the search strategy as indicated by the ratio of studies ultimately included in the review to the number of studies found in the database searches (ie, yield rate)?
4. Did the number of people and time needed to complete a review differ between funded and unfunded reviews (regardless of funding source)?

METHODS
Study selection
We searched the PROSPERO database hosted at the University of York's Centre for Reviews and Dissemination.[8] According to the website, 'PROSPERO is an international database of prospectively registered systematic reviews in health and social care', with an emphasis on intervention studies, although other reviews concerning patient or clinical relevance are also accepted. We retrieved 3684 registered project records from the PROSPERO website on 1 July 2014, and 437 records were marked as having a completed status. Of the 437 records for completed projects, 195 contained a link to one or more publications of a completed review (not simply a protocol). See online supplementary appendix for citations of the 195 publications we analysed. Data were extracted from 195 publications (reporting a total of 197 literature reviews) by two authors (RB, KAK) and were verified by at least one of two additional authors (AWB, PLC). Although an audit trail is available for each PROSPERO record, there is no way to verify the completeness or accuracy of the entries or whether the registrants created the files early in the timeline of the project as prescribed by the registry administrators. Additionally, this registry is not equipped to verify the completeness or accuracy of the entries as compared with the resultant publication. See the text box for clarification of the terms we used in our analysis.

Term used	Data used in Analysis
Authors	The persons listed on the published article reporting the results of the review. Not all team members were listed as authors of resulting publications, so we counted and evaluated 'team members' separately. A sum of unique names was generated to create a combined variable for analysis of people involved (authors/team members).
Team members	The number of persons working on the review per the PROSPERO registry (see above for distinctions with 'authors').
Reviews	The activities of searching the literature, selecting studies that meet inclusion criteria and synthesis of study results. In our data, two governmental publications contained two review processes, thus resulting in our evaluation of 197 reviews reported in 195 publications.
Publications	The published results of the review process(es) in a scholarly journal or government document.
Studies	We counted studies and citations in Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagrams as the same thing, although some levels of PRISMA diagrams may contain some duplicate reports of studies in different citations.

Data extraction process
We extracted the following data from the published reviews: the dates a manuscript was received, accepted and published (received and accepted dates were used only to compare to the registry date to validate the registered timeline); the number of authors; funding information; the data sources searched (total number of databases or websites searched and their names); and the number of studies in the literature filtering steps

from the PRISMA diagrams[9] (when present in the articles). Companion data from the PROSPERO registry for each record included the number of team members, funding, registered start date and geographic location of the team members.

Authors/team members—The authors listed on the publication were counted, as were the team members in the registry. In all but seven cases, the number of authors was larger than the registered number of team members. We added the number of unique authors and registered team members in a combined variable to reflect personnel involvement. We also extracted country and institutional affiliations. Duplicate entries were determined based on comparing titles and team members, with the most complete record being retained.

Time—We estimated the duration of the project as the time from the registered project start date to the publication date of the review. When the article text did not include the needed date information, we sought additional information on the publishers' websites. For articles indexed in PubMed (all but 12), we used the PubMed IDs to extract the dates provided by the publishers to PubMed. These dates were related to the publication timeline (eg, accepted date, first available online, final publication available date) when reported.

Search sources—For the initial total number of studies found in the database searches (highest level of the PRISMA diagram), we recorded what was reported in the PRISMA diagram (or text if the article or supplemental material did not contain a PRISMA diagram). When authors reported searching less commonly used sources of information such as local clinical trial registries or country-specific websites, we counted those sources in the 'other' category and did not include them in the 'total databases searched' variable. Authors do not consistently indicate the same level of the PRISMA diagram for the PRISMA category 'studies identified from other sources', so we aggregated these sources in the variable we called 'total N found' to compare across studies.

Search efficiency—We used yield rate as a metric for search efficiency, calculated by dividing the final number of included studies by the initial number of studies found (excluding duplicates) in the literature search. The final number of included studies was recorded from the PRISMA diagram or the text. If the PRISMA diagram or text separately reported the total number of studies included for quantitative (used in a statistical synthesis) and qualitative (used in a narrative summary only) synthesis, we used the highest number to calculate the overall merged yield rate. For example, if the PRISMA diagram noted that 20 unique studies were included in the review, and 8 were in the quantitative synthesis and 15 were in the qualitative summary, we used 20 for the overall merged yield rate owing to some overlap.

Funding—We coded reviews as being funded if the registry or publication text included explicit statements of review-level funding. We did not code studies as being 'funded' if there were only general statements of salary support or conflict disclosures for individual authors (eg, "Dr Smith is funded by an NIH grant"). Explicit statements of author salary support for the review were recorded separately for analysis.

Statistical analysis

To answer our research questions, we calculated the average number of authors/team members, time from registered start date to publication in weeks and study yield rates. We counted the number of reviews that reported general funding for the project or review-related salary support of authors. We summarised frequency counts for each of the 12 most often reported databases used in literature searches in our sample. All summaries and analyses were calculated with SPSS V.22 (IBM) except where noted. Analysis of variance was used to compare means for time to complete and number of authors/team members between funded and unfunded reviews. The summary literature distillation process was generated with R (R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria; 2012. http://www.R-project.org/; figure 1). Owing to the extreme skewness for the study count variables from the PRISMA diagrams, we calculated z-scores, means, SDs, quartiles and ranges from publications with complete data, removing outliers that were beyond 2.5 SDs to generate figure 1. For reviews with incomplete data for any variable, we summarised only reported data and did not write to authors to request missing information.

RESULTS

Table 1 summarises the results for number of authors/team members, time needed to complete the reviews and yield rate. The mean project length (using the registered project start date to the review's publication date) was 67.3 weeks (SD=31.0; range 6–186 weeks). We also calculated a merged yield rate from the number found without duplicates in the initial search to the final included studies (with some small overlap between quantitative and qualitative studies, n=190). The mean merged yield rate was 2.94% (SD=6.49; range 0–64.71%). Of the countries listed for the locations of registered team members, the UK was most often listed: 62 of the 197 reviews included UK team members.

Figure 1 summarises the literature filtering process across reviews, indicating the initial number of studies found and the interim steps until the final mean number of studies included in the sample we analysed. The use of the 'Büchner' plot (shaped like a Büchner funnel and not to be confused with a 'funnel plot' used to evaluate potential publication bias) allows us to emphasise several key characteristics of the process. First, the Büchner plot allows all data to be plotted in such a way that demonstrates the wide range of included

Figure 1 Aggregated literature filtration process based on counts reported (n=195). Trimmed means are indicated in the boxes and trimmed ranges (±2.5 SDs) are indicated at the right and left of each level. Some reviews were published reporting that zero studies met the criteria for inclusion in the review.
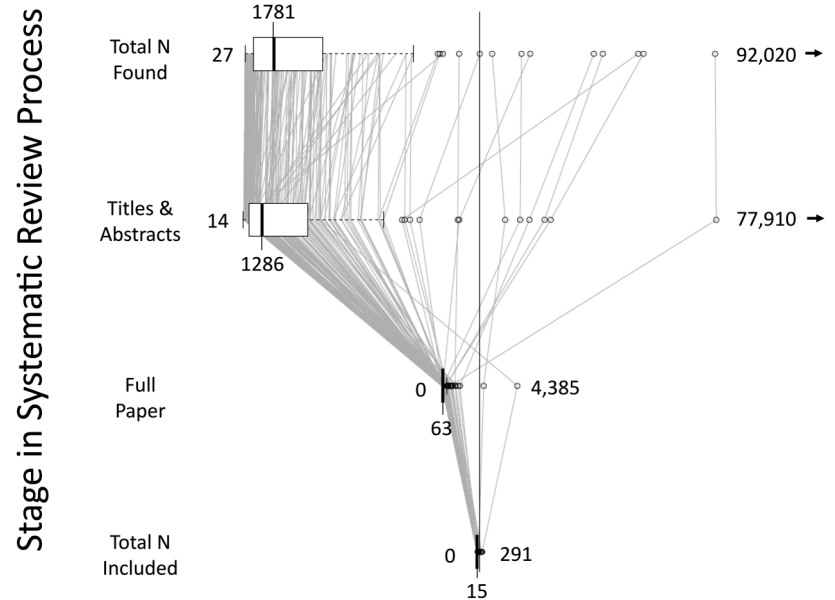
studies at each stage as well as the distribution of the data (highly right-skewed). Second, the ordinality of included studies was not maintained from stage to stage. That is, the reviews with the greatest numbers at one stage were not necessarily the greatest at the next stage. This is visualised by lines crossing throughout the plot. Third, the figure demonstrates that the filtration process can be dramatic, as is reflected in the average yield rate being <3%.

Table 2 summarises the differences in time to complete the review and number of authors/team members per project stratified by reported review level or salary funding. Reviews that reported funding took longer to complete and included more authors/team members; this difference was not seen when we considered only whether specific authors were funded. Table 3 summarises the most commonly used databases reported in the included reviews, with MEDLINE, EMBASE and Cochrane being the top three.

DISCUSSION

Our aim was not to conduct a comprehensive review of all systematic reviews, but rather to generate plausible

estimates of the logistics of conducting reviews for medical interventions for review teams that engaged in at least one best practice: prospective review registration. We therefore chose to evaluate reviews that were registered in the PROSPERO registry and reported to be completed and published. Aspects of performing systematic reviews that remain undocumented (eg, true start date of work, total person hours required), in different domains, or by teams that neither register their projects nor update their registrations on completion may differ. Furthermore, the country affiliation reported in the registry for almost one-third of the team members was the UK. We cannot determine whether our results are representative of all systematic review teams or whether our results reflect an experience more likely to be encountered by UK researchers.

Our convenience sample has several limitations that must be considered. An estimate of the review project start date was difficult to capture from the registry data. The instructions on the PROSPERO site request that registration be done no later than prior to the completion of the data extraction stage.[8] Our calculations for time were anchored by the registered date of the start of the project, but it takes some time to assemble the team,

Table 1 Descriptive statistics for number of authors, time for publication, and quantitative or qualitative yield rates for 195 records analysed in the PROSPERO registry

Category	Mean, SD	Median	IQR	Range
Authors/team members (n=195 publications)	5, 3	5	3	1–27
Time (in weeks; registered project start to publication date, n=192*)	67.3, 31.0	65.8	41.6	6–186
Quantitative analysis yield rate (n=82; %)	2.6, 4.7	1.0	2.7	0.03–32.43
Qualitative analysis yield rate (n=80; %)	2.7, 4.6	1.0	2.5	0.05–26.19
Merged yield rate (n=190; %†)	2.94, 6.49	0.93	2.5	0.0–64.71

*Three studies were excluded from this calculation because they were registered after the publication date.
†Excludes a small overlap between quantitative and qualitative studies when information was not provided in the publication to differentiate between the categories.

Table 2 Comparison (analysis of variance) of reported funding of author salaries (n=20 for both outcomes) or review projects (n=86 for time, n=88 for authors) to time and authors needed to complete and publish the reviews

Outcome	Did outcome differ from reviews reporting no funding?	Means (reported as funded vs funding not reported)	F, p Value
Time to complete (n=191)			df (1, 189)
Review funding reported (n=86)	Yes	42 vs 26 weeks	17.545, <0.001
Salary funding reported (n=20)	No	68 vs 64 weeks	0.258, 0.612
Number of authors/team members (n=195)			df (1193)
Review funding reported (n=88)	Yes	6.8 vs 4.8 persons	14.638, <0.001
Salary funding reported (n=20)	No	6.0 vs 5.7 persons	0.08, 0.778

determine inclusion and exclusion criteria, conduct preliminary searches to refine the search syntax, perform inter-rater reliability for literature screening, and obtain funding (if the review is specifically funded). When a large number of studies are found in a search, these early-stage tasks need to be refined to a standard operating procedure, especially when the evaluation of the initial corpus of found studies cannot feasibly be performed in duplicate. We therefore predict that the time to publication from the very first activities of some reviews may be substantially longer than our estimates. As shown in figure 1, at present, the full text of a large number of papers (median=63, maximum=4385) may need to be screened to evaluate final inclusion criteria. One factor that may impact the timeline is the choice of which and how many databases to search. The varied and mixed sources of searched databases in our data set prevent conclusions about the potential impact of these choices on time and work, but others who have done explicit comparisons have noted that using Web of Science is more efficient that Google Scholar, for example.[10]

Another limitation, as can be seen from figure 1, is that the data at each stage of literature search and selection are very skewed, making attempts to predict

Table 3 Top 12 databases used in included reviews in descending order (N=197 reviews)

Database name	Frequency, per cent
MEDLINE	162, 82.2%
EMBASE	160, 81.2%
Cochrane	148, 75.1%
CINAHL	87, 44.2%
PubMed	58, 29.4%
Web of Science	56, 28.4%
PsycINFO	49, 24.9%
SCOPUS	29, 14.7%
AMED	30, 15.2%
LILACS	24, 12.2%
Google Scholar	13, 6.6%
ProQuest	12, 6.1%

AMED, Allied and Complementary Medicine Database; CINAHL, Cumulative Index to Nursing and Allied Health Literature; LILACS, Literatura Latina-Americana e do Caribe em Ciêcias da Saúde.

timelines from various factors statistically unsound. Other questions may be better answered with data/analyses other than those we reported but using similar approaches. Some examples of other comparisons that could be made with approaches similar to ours include comparisons among types of reviews (eg, interventions vs diagnostic utility), differences between quantitative and qualitative reviews, or whether higher AMSTAR ratings11 are associated with completion time or number of team members. These questions may need to be answered by using systematic surveys or prospective data collection methods.

One study used time logs spent performing 37 reviews to develop a prediction equation on the time to complete a meta-analysis using the number of initial citations retrieved as a predictor,[12] and found that the greatest proportion of time involved was in the preanalysis search, retrieval and database development phase. Overall, using people specialising in this work at a private company, the median total time was 1110 hours, range=216–2518 hours.[12] In contrast, our data most likely reflect the situation in which people who perform systematic reviews may interleave this work with many other job duties, particularly in academia. The average elapsed time in this study is more than 1 year, while the median time reported by Allen and Olkin[12] is a little over a person-year of work time when done by specialists.

Further constraining the present conclusions, the PROSPERO registry does not currently function as a day-to-day project diary or require answers to detailed questions about methods such as the use of automated text screening or data extraction approaches that may affect the timeline or people required. Others who have evaluated the use of text mining[13] and automated data extraction[14] report that these methods may have some utility in certain types of reviews (eg, scoping), but more work is needed to provide significant reductions in work required by humans. Finally, we focused on reviews published in journals rather than those that may be published on websites, and thus the latter may not be subject to delays often encountered in the journal submission and peer-review process unrelated to literature search or synthesis factors. Indeed, the number of

rounds of submission and peer review a particular review went through could add considerably to the time between project initiation and publication.

As the scientific literature continues to grow, generating high-quality, comprehensive reviews in a timely manner will become increasingly costly unless more automated resources are dedicated or different methods to search and retrieve relevant papers can be developed.[4] As noted by the authors of the PRISMA statement, "A systematic review attempts to collate all empirical evidence that fits pre-specified eligibility criteria to answer a specific research question. It uses explicit, systematic methods that are selected with a view to minimizing bias, thus providing reliable findings from which conclusions can be drawn and decisions made."[15] Our results point to an ever-increasing logistical challenge to conducting systematic reviews so that unbiased conclusions can be made about medical interventions. For example, at the transition between the full-text review and the final inclusion stages, our data (shown in figure 1) indicate that a ratio of about 0.76 of papers retrieved are *not* included (untrimmed data ratio=0.68). It cannot be determined from our data how much of this is due to very stringent inclusion/exclusion criteria or incomplete/discrepant reporting, although authors who report the details of why each of the papers were excluded allow readers to know such information for a given review. In some cases, improved reporting and deposition of data in repositories[16] may reduce bias from what otherwise may have resulted in excluded studies.

A recent example that points to the emerging logistical challenges is illustrated in a study that evaluated matched pairs of reviews published on the same research question within 5 years of each other: one using Cochrane review methods and the other from other sources. These authors found a 47% difference in the number of studies in the paired reviews (excluding individual trials included in both reviews) unaccounted for by order of publication.[17] One potential source of search differences may be the use of Medical Subject Headings (MeSH), which were introduced in 1963.[18] Although this system is still able to provide a list of articles on a topic of interest, it is not designed to meet the needs of today's clinicians or systematic reviewers. This approach to indexing and searching the literature, along with author-assigned keywords, is no longer a serviceable approach because of the vast amount of irrelevant or missing articles provided when using such methods, owing to imprecise and overlapping keywords and MeSH terms that provide little context, as well as the lag times between publication and MeSH heading assignment that can occur. An additional potential explanation for the observed differences in reviews reported in this study[17] may be some differences in study selection criteria or databases searched. Clinicians have advised other clinicians who wish to search for studies to inform their clinical practice in PubMed: "Warning: do not look

at all the articles found that although interesting are not pertinent to the present clinical question, or you will be lost in the sea of PubMed!"[19] A search for human clinical trials using the filters provided in PubMed with or without additional index terms can provide results with low sensitivity, precision and specificity.[20] Even something as simple as defining a study as a randomised controlled trial has been observed to be incorrect 20% of the time, which may add non-trivial work to a large review if this leaves many full-text articles to be reviewed before the true study design can be ascertained.[21] CONSORT reporting guidelines have improved this problem by adding study design as a requirement in the title, but even top-tier journals still do not enforce compliance 100% of the time.[22]

Systematic reviews and meta-analyses can require large amounts of time and effort to complete, with often (based on our own experiences) unpredictable uncertainty in the time and resources required to complete a review. Although statistical methods have been developed to estimate trends in large corpora of literature,[21] and crowdsourcing may help to decrease the calendar time needed to complete a review,[23] we assert that the time is ripe to investigate metadata[4 24] approaches to indexing publications so that more targeted yet comprehensive searches can be performed efficiently with high specificity and precision. We suggest that human clinical trials could be a starting point for testing the use of indexing systems that employ a Population, Intervention, Comparisons, Outcomes, Study Design (PICOS)[15] structure of coding metadata for clinical trials in order to make the scientific literature become a truly searchable database. Our preliminary testing has shown that a modestly trained worker can code a single paper in about 15 min. By application of the recently proposed Findable, Accessible, Interoperable and Reusable (FAIR) metadata guidelines,[4] clinical trial data (or other types of research data) could be encoded on publication so that the data are in fact findable, accessible and thus potentially more interoperable and reusable. An added benefit of this approach is that it would make the recently proposed mandated data sharing[2] essentially automatic, leaving the critical appraisal, synthesis and interpretation to be done. In a brighter future, evidence-based, up-to-date summaries could be produced on demand with less human effort and may reduce the delay between the question and the evidence-based answer.

REFERENCES

1. Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLoS Med* 2010;7:e1000326.
2. European Science Foundation. Forward Look—Implementation of Medical Research in Clinical Practice. European Science Foundation, 2011. http://www.esf.org/fileadmin/Public_documents/Publications/Implem_MedReseach_ClinPractice.pdf (accessed 1 Apr 2016).
3. Taichman DB, Backus J, Baethge C, et al. Sharing clinical trial data: a proposal from the international committee of medical journal editors. *Ann Intern Med* 2016;164:505–6.
4. Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 2016;3:160018.
5. The Cochrane Collaboration. Cochrane Database of Systematic Reviews. 2015. https://community.cochrane.org/editorial-and-publishing-policy-resource/cochrane-database-systematic-reviews-cdsr (accessed 1 Apr 2016).
6. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. http://handbook.cochrane.org/
7. Stovold E, Beecher D, Foxlee R, et al. Study flow diagrams in Cochrane systematic review updates: an adapted PRISMA flow diagram. *Syst Rev* 2014;3:54.
8. PROSPERO—International Prospective Register for Systematic Reviews. (University of York, Centre for Reviews and Dissemination, United Kingdom, 2015). http://www.crd.york.ac.uk/prospero/ (accessed 1 Apr 2016).
9. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;338:b2535.
10. Levay P, Ainsworth N, Kettle R, et al. Identifying evidence for public health guidance: a comparison of citation searching with Web of Science and Google Scholar. *Res Synth Methods* 2016;7:34–45.
11. Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol* 2009;62:1013–20.
12. Allen IE, Olkin I. Estimating time to conduct a meta-analysis from number of citations retrieved. *JAMA* 1999;282:634–5.
13. Shemilt I, Simon A, Hollands GJ, et al. Pinpointing needles in giant haystacks: use of text mining to reduce impractical screening workload in extremely large scoping reviews. *Res Synth Methods* 2014;5:31–49.
14. Jonnalagadda SR, Goyal P, Huffman MD. Automating data extraction in systematic reviews: a systematic review. *Syst Rev* 2015;4:78.
15. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009;6:e1000100.
16. Marcial LH, Hemminger BM. Scientific data repositories on the Web: an initial survey. *J Am Soc Inform Sci Technol* 2010;61:2029–48.
17. Useem J, Brennan A, LaValley M, et al. Systematic differences between Cochrane and non-Cochrane meta-analyses on the same topic: a matched pair analysis. *PLoS ONE* 2015;10:e0144980.
18. Rogers FB. Communications to the editor. *Bull Med Libr Assoc* 1963;51:114–16.
19. Costantino G, Montano N, Casazza G. When should we change our clinical practice based on the results of a clinical study? Searching for evidence: PICOS and PubMed. *Intern Emerg Med* 2015;10:525–7.
20. Hoogendam A, de Vries Robbe PF, Stalenhoef AF, et al. Evaluation of PubMed filters used for evidence-based searching: validation using relative recall. *J Med Libr Assoc* 2009;97:186–93.
21. Capers PL, Brown AW, Dawson JA, et al. Double sampling with multiple imputation to answer large sample meta-research questions: introduction and illustration by evaluating adherence to two simple CONSORT guidelines. *Front Nutr* 2015;2:6.
22. Hays M, Andrews M, Wilson R, et al. Reporting quality of randomised controlled trial abstracts among high-impact general medical journals: a review and analysis. *BMJ Open* 2016;6:e011082.
23. Brown AW, Allison DB. Using crowdsourcing to evaluate published scientific literature: methods and example. *PLoS ONE* 2014;9:e100647.
24. National Information Standards Organization. Understanding Metadata. 2004. http://www.niso.org/standards/resources/UnderstandingMetadata.pdf. (accessed 19 Jan 2016).