This is a repository copy of *A third of systematic reviews changed or did not specify the primary outcome:A PROSPERO register study*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/98989/

Version: Accepted Version

Article:
Tricco, Andrea C, Cogo, Elise, Page, Matthew J et al. (8 more authors) (2016) A third of systematic reviews changed or did not specify the primary outcome:A PROSPERO register study. Journal of Clinical Epidemiology. 46–54. ISSN: 0895-4356

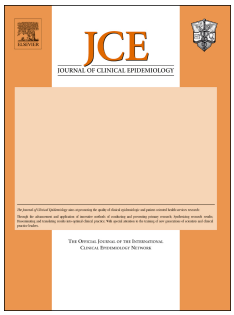
https://doi.org/10.1016/j.jclinepi.2016.03.025

Reuse
This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Accepted Manuscript



A third of systematic reviews changed or did not specify the primary outcome: A PROSPERO register study

Andrea C. Tricco, Elise Cogo, Matthew J. Page, Julie Polisena, Alison Booth, Kerry Dwan, Heather MacDonald, Tammy J. Clifford, Lesley A. Stewart, Sharon E. Straus, David Moher

PII: S0895-4356(16)30076-2

DOI: 10.1016/j.jclinepi.2016.03.025

Reference: JCE 9156

To appear in: *Journal of Clinical Epidemiology*

Received Date: 9 April 2015

Revised Date: 8 March 2016

Accepted Date: 10 March 2016

Please cite this article as: Tricco AC, Cogo E, Page MJ, Polisena J, Booth A, Dwan K, MacDonald H, Clifford TJ, Stewart LA, Straus SE, Moher D, A third of systematic reviews changed or did not specify the primary outcome: A PROSPERO register study, *Journal of Clinical Epidemiology* (2016), doi: 10.1016/j.jclinepi.2016.03.025.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

A third of systematic reviews changed or did not specify the primary outcome: A PROSPERO register study

Andrea C Tricco[a,b] email: TriccoA@smh.ca

Elise Cogo[a] email: CogoE@smh.ca

Matthew J Page[c] email: Matthew.page@monash.edu

Julie Polisena[d,e] email: Juliep@cadth.ca

Alison Booth[f] email: Alison.booth@york.ac.uk

Kerry Dwan[g] email: Kerry.dwan@liverpool.ac.uk

Heather MacDonald[a] email: Hrmacdonald@gmail.com

Tammy J Clifford[d] email: TammyC@cadth.ca

Lesley A Stewart[f] email: Lesley.stewart@york.ac.uk

Sharon E Straus [a,h] email: StrausS@smh.ca

David Moher[i] email: Dmoher@ohri.ca

[a]Knowledge Translation Program, Li Ka Shing Knowledge Institute, St. Michael's Hospital, 209 Victoria Street, East Building. Toronto, Ontario, M5B 1T8, Canada

[b]Epidemiology Division, Dalla Lana School of Public Health, University of Toronto, 155 College Street, 6th floor, Toronto, Ontario M5T 3M7, Canada

[c]School of Public Health & Preventive Medicine, Monash University, Level 6, The Alfred Centre, 99 Commercial Road, Melbourne, Victoria 3004, Australia

[d]Canadian Agency for Drugs and Technologies in Health (CADTH), 865 Carling Avenue Suite 600, Ottawa, Ontario K1S 5S8

[e]Department of Epidemiology and Community Medicine, University of Ottawa, 451 Smyth Road, Ottawa, Ontario, K1H 8M5, Canada

24 [f]Centre for Reviews and Dissemination, University of York, York YO10 5DD, UK

25 [g]Department of Biostatistics, University of Liverpool, UK

26 [h]Department of Geriatric Medicine, University of Toronto, 27 Kings College Circle. Toronto,

27 Ontario M5S 1A1, Canada

28 [i]Ottawa Methods Centre, Ottawa Hospital Research Institute, The Ottawa Hospital, 501 Smyth

29 Road, PO BOX 201B, Ottawa, Ontario, Canada, K1H 8L6

30 **Correspondence to:**

31 Andrea Tricco, MSc, PhD

32 Scientist, Knowledge Translation Program

33 Li Ka Shing Knowledge Institute, St. Michael's Hospital

34 209 Victoria Street, East Building, Room 716, Toronto, Ontario, M5B 1W8, Canada

35 Email: triccoa@smh.ca, Telephone: 416-864-6060 x77521

ACCEPTED MANUSCRIPT

ABSTRACT

Objectives: To examine outcome reporting bias of systematic reviews registered in PROSPERO.

Study Design and Setting: Retrospective cohort study. The primary outcomes from systematic review publications were compared with those reported in the corresponding PROSPERO records; discrepancies in the primary outcomes were assessed as upgrades, additions, omissions or downgrades. Relative risks (RR) and 95% confidence intervals (CI) were calculated to determine the likelihood of having a change in primary outcome when the meta-analysis result was favourable and statistically significant.

Results: 96 systematic reviews were published. A discrepancy in the primary outcome occurred in 32% of the included reviews and 39% of the reviews did not explicitly specify a primary outcome(s); 6% of the primary outcomes were omitted. There was no significant increased risk of adding/upgrading (RR 2.14, 95% CI 0.53 to 8.63) or decreased risk of downgrading (RR 0.76, 0.27-2.17) an outcome when the meta-analysis result was favourable and statistically significant. As well, there was no significant increased risk of adding/upgrading (RR 0.89, 0.31-2.53) or decreased risk of downgrading (RR 0.56, 0.29-1.08) an outcome when the conclusion was positive.

Conclusions: We recommend review authors carefully consider primary outcome selection and journals are encouraged to focus acceptance on registered systematic reviews.

Word count: 200 (abstract), 3286 (main text), 2 figures, 3 tables, 14 appendices.

Keywords: bias, methodology, quality, reporting, systematic reviews, outcome reporting bias

ACCEPTED MANUSCRIPT

57 **Running title:** Examining outcome reporting bias in systematic reviews

ACCEPTED MANUSCRIPT

What is new?

Key finding

- Many systematic reviews that are registered in PROSPERO have discrepancies in primary outcomes between their record and review publication.

What this study adds to what is known?

- This is the first study to examine outcome reporting bias using the PROSPERO register, a database for prospectively registering systematic reviews that was established in 2011.
- Previous studies have compared outcomes reported in Cochrane reviews to those reported in the corresponding review protocols. These studies found that more than 1/3 of published systematic reviews had a discrepancy between the outcomes reported in the protocol versus final publication. One study found evidence of outcome reporting bias, in which statistically significant outcomes were more likely to be upgraded (i.e. promoted from secondary to primary) or added in the final publication compared to the protocol.
- We found that approximately 1/3 of published systematic reviews had a discrepancy between the outcomes reported in the PROSPERO record versus the review publication. However, evidence of outcome reporting bias was not observed.

What is the implication, and what should change now?

- Our study suggests that non-Cochrane review authors have similar outcome reporting behaviours to Cochrane review authors. We recommend that all non-Cochrane reviews are registered with PROSPERO, review authors carefully consider the selection of primary outcomes, peer reviewers should check PROSPERO to see if there are any discrepancies between the record and review publication, and journals are encouraged to focus acceptance on registered systematic reviews.

ACCEPTED MANUSCRIPT

1. INTRODUCTION

The Cochrane Handbook for Systematic Reviews of Interventions [1] states that systematic reviewers should prepare a systematic review protocol prior to their review conduct, to encourage transparency of reporting hypotheses and methods (including outcomes) and avoid outcome reporting bias. This is consistent with the Institute of Medicine Standards for Systematic Reviews [2]. As well, the Cochrane Handbook [1] and Preferred Reporting Items for Systematic reviews and Meta-Analyses Statement [3] state that any changes to the protocol should be fully documented and explained in the systematic review publication. Despite this guidance, research consistently has found that more than 1/3 of published systematic reviews have an undisclosed discrepancy between the outcomes reported in the protocol versus final review [4-7].

In the most simplistic definition, outcome reporting bias "occurs when a study in which multiple outcomes were measured reports only those that are [statistically] significant" [8]. Previous studies have compared final Cochrane review methods to those reported in the review protocols [4-7], including a recent Cochrane methodology review on outcome reporting bias [9]. One of these studies found evidence of outcome reporting bias, in which statistically significant outcomes were more likely to be upgraded (i.e. promoted from secondary to primary) or added in the final publication compared to the protocol [5]. All of these studies included a sample of systematic reviews published in the Cochrane Database of Systematic Reviews prior to the year 2009.

The International Prospective Register of Systematic Reviews (PROSPERO) was established in 2011 [10] and is the only open access online facility to prospectively register non-Cochrane systematic reviews. Since most published systematic reviews are not Cochrane reviews

81 [11], this register of review protocol details is likely a more representative sample of systematic
82 reviews in the literature. No previous study has explored outcome reporting bias of systematic
83 reviews registered in PROSPERO. As such, we aimed to 1) examine whether outcome reporting
84 bias exists, and to what extent, in published systematic reviews registered in PROSPERO, as
85 well as 2) assess the methodological quality of published systematic reviews that were registered
86 in PROSPERO.

ACCEPTED MANUSCRIPT

2. METHODS

2.1 Protocol

Prior to conducting this retrospective cohort study, we created a project plan, which outlined our study methods. Our protocol was revised after receiving feedback from all authors. The final protocol can be found in Appendix A. Since this study was not a systematic review, it was not eligible to be registered with the PROSPERO repository.

2.2 Sample of systematic reviews

We aimed to identify all completed systematic reviews of interventions that were registered in PROSPERO. On November 29, 2013, all records from the PROSPERO database identified as "Completed and Published" were downloaded. These records also include the citation/link to the final publication. PROSPERO includes an audit trail for protocol amendments and progress reports. For the purpose of our study, the protocol record used was the version immediately prior to the version where the Named Contact updated the record to report that the review had been completed. Our scope was limited to systematic reviews of interventions to allow the comparison of statistically significant meta-analysis results, which would not be feasible for other review products (e.g., diagnostic reviews, prognostic reviews, prevalence reviews). Only non-Cochrane reviews were included. Completed reviews not published in English were also excluded, due to resource limitations.

2.3 Data abstraction process

A data abstraction form with an explanation guide was developed (Appendix Table A) and calibrated through a team exercise. Specifically, the team independently pilot-tested the forms using a random sample of 10 included systematic reviews. Data abstraction did not commence until high agreement (>90%) was achieved. Subsequently, 3 pairs of reviewers abstracted each of the systematic review publications, independently. In order to ensure

consistency across the team regarding the classification of outcomes, one team member verified all of the data (EC) and resolved discrepancies.

2.4 Data items

The data items were abstracted from both the protocol details and the publication, and included study characteristics (e.g., year of publication, number of studies included, type of studies included, whether meta-analysis was conducted, source of funding), number of primary outcomes, changes in primary outcomes from the PROSPERO record to review publication, reasons for changes in primary outcomes (if reported), meta-analysis results, and conclusions. The reason we focused on primary outcomes is because this is the outcome of greatest interest and importance. Similar research on outcome reporting bias has used this approach [4-7].

If the primary outcome(s) was not explicitly stated in the publication (i.e. not specifically called a "primary" outcome), the following decision-tree approach [12, 13] was used to "derive" the primary outcome(s), by selecting the outcome that met the first of the following criteria: (1) the outcome(s) listed in the title; (2) the outcome(s) listed in the objectives; (3) the most serious outcome (e.g., mortality). To facilitate comparison across studies, all changes in primary outcomes from the PROSPERO record to the systematic review publication were coded using the same classification scheme used in the Parmelli *et al*. [7] and Kirkham *et al.* [5] studies. These categories were new inclusion of outcomes (or additions), exclusion, upgrade, and downgrade of outcomes (Box). The meta-analysis results were categorized using a previous approach [13], including favourable and statistically significant, favourable and not statistically significant, neutral, unfavourable and not statistically significant, and unfavourable and statistically significant (Box, Appendix Figure A). The conclusions were obtained from the

abstract and discussion sections from the systematic reviews and were categorized using a previous approach [13], including positive, neutral, negative, and indeterminate (Box).

We used the same hierarchy reported by Kirkham *et al.* to select meta-analyses from systematic reviews with multiple treatment group comparisons [5]. Specifically, we selected the first intervention comparison which met the following criteria: "(1) an intervention comparison described in the protocol as the primary review comparison; (2) the first intervention comparison mentioned in the title of the protocol; (3) an intervention comparison described in the review as the primary review comparison; (4) the first intervention comparison mentioned in the objectives of the review; (5) the intervention comparison used in the first meta-analysis presented in the review."

2.5Methodological quality appraisal

The overall methodological quality of the systematic reviews was assessed using the Assessment of Multiple SysTemAtic Reviews (AMSTAR) tool (Appendix Table B) [14]. The scores range from 0 to 11, with higher scores indicating superior quality. For our study, a score of 8 or higher was considered higher quality. This assessment was conducted to ascertain the overall quality of completed and published systematic reviews that were registered in PROSPERO.

2.6 Analysis

We explored the association between statistical significance of meta-analysis results and adding, upgrading or downgrading of outcomes compared to no discrepancies, by calculating a relative risk (RR) and 95% confidence interval (CI), where the meta-analysis results were dichotomised into favourable and statistically significant versus any of the other 4 categories. The formula is RR=[a/(a+b)]÷[c/(c+d)], where a is the number of meta-analysis outcomes that

are discrepant and have a favourable and statistically significant result, b is the number of meta-analysis outcomes that are not discrepant and have a favourable and statistically significant result, c is the number of meta-analysis outcomes that are discrepant and do not have a favourable and statistically significant result, and d is the number of meta-analysis outcomes that are not discrepant and do not have favourable and statistically significant result. This analysis was similar to those conducted by Page and colleagues in their Cochrane review of outcome reporting bias [9]. The RR and 95% CI were calculated for outcomes that were explicitly reported as primary outcomes, as well as including those that were derived using the classification scheme reported above. Our hypotheses were that when the meta-analysis result was favourable and statistically significant, adding/upgrading of outcomes would be more likely while downgrading of outcomes would be less likely. A sensitivity analysis was also conducted consistent with the analysis method used by Kirkham and colleagues [5], to allow comparability of results. For this analysis, the meta-analysis results were dichotomised into statistically significant versus not statistically significant and the hypotheses were that new/upgraded outcomes would be more likely to have statistically significant meta-analysis results while downgraded outcomes would be less likely, than if there was no discrepancy.

We also conducted a post-hoc analysis for systematic reviews that were funded. Similar to our primary analysis, we explored the association between statistical significance of meta-analysis results and adding, upgrading or downgrading of outcomes compared to no discrepancies by calculating a RR and 95% CI, where the meta-analysis results were dichotomised into favourable and statistically significant versus any of the other 4 categories. This analysis was repeated for systematic reviews that did not have funding. Sensitivity analyses were also conducted using the Kirkham *et al.* approach [5].

The RR and 95% CI were calculated for obtaining a positive conclusion for new primary outcomes or upgrades, and downgrades compared to no discrepancies (where conclusions were categorised as positive versus all other conclusion types). Our hypotheses were that when the conclusion was positive, adding/upgrading of outcomes would be more likely while downgrading of outcomes would be less likely. A sensitivity analysis was also conducted to calculate the RR and 95% CI using a similar approach as to Kirkham *et al.* [5]. For this sensitivity analysis, our hypothesis was that when outcomes were added or upgraded, a positive conclusion would be more likely, while when outcomes were downgraded, a positive conclusion would be less likely.

ACCEPTED MANUSCRIPT

3. RESULTS

3.1 Sample of PROSPERO records

In November 2013, 2,426 protocol records were registered with PROSPERO and 344 were completed systematic reviews (Figure 1). Of the completed reviews, 140 were potentially relevant (i.e., published or *in press*), and of these 44 were excluded because they were not systematic reviews of interventions or the final review was not written in English (Appendix Table C). Ninety-six systematic reviews fulfilled the eligibility criteria and were subsequently included (Appendix Table C).

3.2 Systematic review characteristics

Eighty-nine (92.7%) of the systematic reviews were published between 2012 and 2013, and 4 (4.2%) were published in 2014, as they were *in press* at the time we downloaded their PROSPERO records. 81 (84.3%) included 2 to 30 studies, 56 (58.3%) limited inclusion to randomized controlled trials, and 67 (68.8%) conducted a meta-analysis (Table 1). In addition, 36 (37.5%) reported no source of funding, 45 (46.9%) were conducted in the United Kingdom or North America, and 5 (5.2%) published their protocol in a journal.

3.3 Methodological quality

Eight of the 11 AMSTAR items were adequately addressed by more than 72 (75%) of the systematic reviews (Figure 2, Appendix Table D). However, 72 (75%) of the reviews did not state conflicts of interest for included studies and review authors, 63 (66%) did not provide a list of excluded studies, 39 (41%) did not assess publication bias where it would have been appropriate to do so, and 14 (15%) did not consider methodological quality or risk of bias results in their conclusion statements.

3.4 Outcome reporting

Although the primary outcome was indicated in PROSPERO, which is structured to separate primary and secondary outcomes, it was not explicitly reported for 37 (38.5%) of the completed systematic reviews, so was derived for the purpose of our study (Table 2). The primary outcomes were derived using the title (35.2%), objectives (24.3%), or were the most serious outcomes (40.5%). Thirty-one (32.3%) of the systematic reviews had a discrepancy between the primary outcomes reported in the PROSPERO record and final publication, while 65 (67.7%) had no discrepancies (Table 3). Of the reviews with discrepancies, 6 (5.9%) had a new primary outcome, 6 (5.9%) excluded a primary outcome, 6 (5.9%) upgraded an outcome, and 22 (21.8%) downgraded a primary outcome. One (1.0%) of the systematic reviews reported a reason for changing their primary outcome. Six (5.9%) systematic reviews reported a change in their primary outcome definition and 1 (1.0%) changed the measurement method for the primary outcome.

3.5 Meta-analysis results

The results of 139 meta-analyses in 67 systematic reviews are presented in Appendix Table E. There was no significant increased risk of adding or upgrading an outcome when the meta-analysis result was favourable and statistically significant (RR 2.14, 95% CI 0.53 to 8.63), which was the same result as found in our sensitivity analysis (Appendix Table F). This result was unchanged when only the primary outcomes that were explicitly reported were included in our analysis (RR 2.02, 95% CI 0.35 to 11.56; Appendix Table G). Further, there was no significant decreased risk of downgrading an outcome when the meta-analysis result was favourable and statistically significant (RR 0.76, 95% CI 0.27 to 2.17) and the same result was observed in our sensitivity analysis. Similarly, when only the primary outcomes that were explicitly reported were included in our analysis, no statistically significant results were

observed for downgrades (RR 1.37, 95% CI 0.20 to 9.42). Calculations were not possible for excluded primary outcomes since they were absent from the publications (by definition).

A post-hoc analysis was conducted for systematic reviews with funding, as well as for systematic reviews without funding (Appendix Tables H-J). No statistically significant results were observed in our overall analysis or sensitivity analyses.

3.6 Conclusion statements

The categorisation of conclusions for all included systematic reviews is presented in Appendix Table K. There was no significant increased risk of adding or upgrading outcomes when the conclusion was positive (RR 0.89, 95% CI 0.31 to 2.53). Further, there was no significant decreased risk of downgrading an outcome when the conclusion was positive (RR 0.56, 95% CI 0.29 to 1.08). Our sensitivity analyses also found no significant risk of a positive conclusion when the outcomes were added/upgraded or downgraded (Appendix Table L).

ACCEPTED MANUSCRIPT

4. DISCUSSION

One-third of published systematic reviews that were registered with PROSPERO had a discrepancy between the primary outcome reported in their record and the primary outcome reported in the review publication. Of the discrepancies, downgrading of primary outcomes was most common (22%), and 6% of reviews omitted a protocol-specified primary outcome from the review. In addition, 39% of reviews did not explicitly specify a primary outcome(s) in the review. Although a lot of discrepancies were observed, we did not find statistically significant associations between discrepant outcome reporting and having a favourable and statistically significant meta-analysis result or positive conclusion. However, the small number of reviews within each subgroup of discrepancy classification likely limited the statistical power to detect statistically significant results. PROSPERO has now passed 5,000 registrants and repeating this study is likely to yield a larger number of published systematic reviews to examine.

Our study is the first to measure outcome reporting bias of systematic reviews that were registered in PROSPERO. To examine this issue, we systematically searched for 96 systematic reviews published between 2011 and 2014. We abstracted data in duplicate, which were triple-checked by a third reviewer, and appraised the included reviews using the AMSTAR tool. The included systematic reviews were of high methodological quality, on average. Areas for improvement included providing a list of excluded studies, assessing publication bias when appropriate (as per the AMSTAR criterion), and reporting conflicts of interest for the systematic review authors, as well as for the included studies.

Our results are only generalizable to intervention reviews, as the risk of outcome reporting bias in other types of reviews (e.g., diagnostic reviews) remains unknown. As well, we only included non-Cochrane reviews. We considered only primary outcomes, which may have underestimated the occurrence of outcome reporting bias for all types of outcomes. However,

this is the same approach to other studies examining outcome reporting bias [4-7]. Limited resources meant that we were unable to contact authors of the discrepant systematic reviews to determine the reason for these inconsistencies. Only one review reported a rationale for changing the outcome, which makes it difficult to provide definitive conclusions as to why these changes may occur [15]. The reason that was reported by the authors was that the clinical experts on their team selected the most clinically important outcomes, which did not align with what was reported in their PROSPERO record. We were unable to include a larger sample of published and completed systematic reviews, due to resource restraints. Due to the small number of included reviews in our analyses, we were unable to examine possible sources of heterogeneity that may have confounded our results or conduct sub-group analysis for outcome reporting bias for systematic reviews with active comparators versus placebo, “high” versus “low” quality as per the AMSTAR tool, and randomized trials versus non-randomized studies. As well, there is a chance that there were more completed systematic reviews that were published but the authors of the review failed to update their PROSPERO record (although they are sent 3auto- reminders to update their information in PROSPERO). We were only able to include the systematic reviews with meta-analyses in our statistical analysis of outcome reporting bias, which is consistent with previous studies [4-7]. Finally, we calculated risk ratios instead of odds ratios to compare our study with previous studies conducted in this area.

A recent Cochrane review [9] included 4 previous studies that examined discrepancies in outcome reporting between systematic review protocols and published systematic reviews [4-7]. All of these studies included Cochrane reviews that were published between 2000 and 2009 and none appraised the methodological quality of included systematic reviews using the AMSTAR tool. A total of 485 Cochrane Reviews were included and discrepancies were identified in 38%

of these. A meta-analysis of two of the studies was conducted and no statistically significant association between statistical significance of meta-analysis results and discrepant outcome reporting (adding, upgrading or downgrading) was found. These results are consistent with those observed in our study.

Our results suggest that authors of non-Cochrane reviews are similar to Cochrane review authors in their outcome reporting behaviours. It is possible that systematic review authors are not focused on identifying primary outcomes of interest at the protocol stage, and are instead just completing the PROSPERO form. Further, as registration in PROSPERO is voluntary (and is relatively new) it is possible that our sample (as well as studies using samples of Cochrane reviews) underestimated the overall number of primary outcome discrepancies in systematic reviews in general.

Using pre-established methods [16], we estimate that 17,399 systematic reviews were published in 2013. During this time, 1,612 Cochrane reviews were registered and 1,526 non-Cochrane reviews were registered with PROSPERO. This means that only 18% of published systematic review authors registered their protocol. As such, we recommend that all non-Cochrane reviews are registered with PROSPERO. Furthermore, review authors are advised to consider the selection of primary outcomes carefully and report the explanations for protocol modifications in the final review publication. Review authors should think about the importance of outcomes prior to embarking on their review and limit the number of outcomes to ensure that those selected are both necessary and meaningful. Core outcome sets have been recommended for trials (COMET initiative, http://www.comet-initiative.org/) and it is recommended that systematic review authors are familiar with this guidance when selecting outcomes for inclusion in their review. Peer reviewers should check PROSPERO to see if there are any discrepancies

between the record and review publication and ensure that the author explains these. Finally, journals are encouraged to focus acceptance on registered systematic reviews, as we found that these are likely to be of high methodological quality.

Few studies have examined outcome reporting bias in systematic reviews [9]. There has been no study of systematic reviews that are not registered with the Cochrane Collaboration or PROSPERO. This could be done by contacting review authors to obtain their unpublished protocol, if one exists. Future research should examine a larger sample of PROSPERO records as this database matures, as well as examine the discrepancies in primary outcomes reported in the abstract and full-text of the published systematic reviews.

ACCEPTED MANUSCRIPT

AUTHORS' CONTRIBUTIONS

All authors conceptualized the study. ACT pilot-tested the data abstraction form, resolved discrepancies, analyzed the results, interpreted the results, wrote the paper, and approved the final paper. EC coordinated the review, pilot-tested the data abstraction form, resolved discrepancies, checked all of the cleaned data, helped write the paper, and approved the final paper. AB, JP, TC, KD, MJP, and HM pilot-tested the data abstraction form, conducted data abstraction, appraised the quality of the articles, edited the paper, and approved the final paper. MJP also analyzed the data, AB screened the records for inclusion, and HM helped clean the data and resolve discrepancies. TC, LAS, SES, and DM edited the paper and approved the final paper.

ACT accepts full responsibility for the finished article, had access to all of the data, and controlled the decision to publish. ACT affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that no discrepancies from the study as planned occurred.

ACKNOWLEDGEMENTS

We thank Inthuja Selvaratnam for formatting our paper and Judy Tran for assisting with creating the manuscript tables. ACT affirms everyone who contributed significantly to the work in this study has been acknowledged.

COMPETING INTERESTS

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no financial support for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous three years; AB, LAS, and DM are members of the PROSPERO Advisory Group; ACT is an author of one of the included systematic reviews but was not involved with the

ACCEPTED MANUSCRIPT

AMSTAR appraisal or data abstraction for this review and was blinded to the author names during the analysis, she is also an Associate Editor for the journal but was not involved with the decision to publish; no other relationships or activities that could appear to have influenced the submitted work.

FUNDING

No funding was received for this study. ACT is funded by a Canadian Institutes of Health Research New Investigator Award in Knowledge Synthesis, SES is funded by a Tier 1 Canada Research Chair in Knowledge Translation, AB and LAS are funded by the National Institute for Health Research, and DM is funded by a University of Ottawa Research Chair.

ETHICS APPROVAL

Ethics approval was not required.

COPYRIGHT/LICENSE FOR PUBLICATION

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, a worldwide license to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the Contribution, v) the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, vi) license any third party to do any or all of the above.

Box 1 Classification: Primary outcomes, Meta-analysis results, and Conclusion statements

Classification of changes to primary outcomes:

- New (Inclusion or Addition): the addition of a completely new primary outcome;
- Exclusion: the omission of a primary outcome in the publication;
- Upgrade: when a secondary outcome in the protocol was changed to a primary outcome in the publication;
- Downgrade: when a primary outcome in the protocol was changed to a secondary or undefined outcome in the publication.

Classification of meta-analysis results:

- Favourable, statistically significant (i.e. effect in favour of the intervention with $p \leq 0.05$);
- Favourable, non-statistically significant;
- Neutral (effect size between 0.95-1.05 and the confidence interval crosses 1);
- Unfavourable, statistically significant (i.e. effect in favour of the non-intervention comparator with $p \leq 0.05$);
- Unfavourable, non-statistically significant.

Categorization of conclusion statements

- Positive (authors stated that there is evidence of effectiveness);
- Neutral (no evidence of effectiveness or they reported no opinion);
- Negative (authors advised against the use of the intervention or it was not recommended); or
- Indeterminate (authors stated that there is insufficient evidence or that more research is required).

369

ACCEPTED MANUSCRIPT

370 **FIGURE LEGENDS**

371 **Figure 1: Flow of systematic reviews through the study**

372 **Figure 2: AMSTAR methodological quality results**

373 Note: NA = not applicable.

374 Items:

375 1. A priori design

376 2. Duplicate selection/DA

377 3. Literature search

378 4. Publication status

379 5. List of studies

380 6. Study characteristics

381 7. Quality assessed

382 8. Quality used

383 9. Methods appropriate

384 10. Publication bias assessed

385 11. Conflicts stated

ACCEPTED MANUSCRIPT

386 **Table 1. Characteristics of the 96 included systematic reviews**

Characteristic	# of systematic reviews (%)
Publication year	
2011	3 (3.1)
2012	29 (30.2)
2013	60 (62.5)
2014	4 (4.2)
Total # of studies included	
0-20	70 (72.9)
21-40	9 (19.8)
>40	7 (7.3)
Total # of participants in included studies	
≤1000 to 5000	48 (50)
5001-10,000	5 (5.2)
10,001-50,000	7 (7.3)
50,001-100,000	3 (3.1)
>100,000	2 (2.1)
Not Reported	31 (32.3)
Study designs included	
All randomized controlled trials	56 (58.3)
Mixed study designs*	35 (36.5)
All observational studies	5 (5.2)
Meta-analysis conducted	
Yes	67 (69.8)
No	29 (30.2)
Funding†	
Stated no funding received	36 (37.5)
Public funder (e.g., academia, government)	56 (58.4)
Commercial Organization	4 (4.2)
Geographic Region‡	
Europe	47 (49)
North America	20 (20.9)
South America	11 (11.4)
Easter Asia	9 (9.3)
Australia	5 (5.2)
Southern Asia	2 (2.1)
Southern Africa	2 (2.1)
Published protocol in a journal	
Yes	5 (5.2)
No	91 (94.8)
Participant population in publication§	
Healthy or presumed healthy	14 (14.6)
Mixed conditions	11 (11.5)
Musculoskeletal conditions	10 (10.4)
Infectious diseases	9 (9.4)
Present/history of cancer	9 (9.4)
Pregnancy-related or reproductive conditions	8 (8.3)
Psychiatric/mental health conditions	7 (7.3)
Cardiovascular conditions	6 (6.3)
Respiratory conditions	6 (6.3)
Autoimmune diseases	3 (3.1)
Gastrointestinal and abdominal conditions	2 (2.1)
Genetic diseases	2 (2.1)
Neurodegenerative/neurological conditions	2 (2.1)
Oral-related conditions	2 (2.1)
Urinary conditions	2 (2.1)
Auditory conditions	1 (1.0)
Overweight	1 (1.0)
Type 2 diabetes	1 (1.0)

ACCEPTED MANUSCRIPT

Note: *Mixed could indicate, for example, RCT & quasi-RCT (not necessarily mixed with observational studies); † Source: Cochrane EPOC Group. Available at: http://epoc.cochrane.org/sites/epoc.cochrane.org/files/uploads/datacollectionchecklist.pdf; ‡ If more than one country was listed (n = 8), only the first country's geographic region is listed here; §as reported by the review authors.

ACCEPTED MANUSCRIPT

392 **Table 2. Number of Primary Outcomes in the Publications**

Outcome details	# of systematic reviews (%)
Number explicit per review	
0	37 (38.5)
1	35 (36.5)
2	10 (10.4)
3	6 (6.3)
4	3 (3.1)
5	1(1.0)
6	2 (2.1)
7	1 (1.0)
8	1 (1.0)
Number derived per review	
NA (were explicit)	59 (61.5)
1	24 (25.0)
2	6 (6.3)
3	5 (5.2)
4	1 (1.0)
5	0 (0)
6	1 (1.0)
Derived Method Used	
NA (were explicit)	59 (61.5)
Method 1-from title	13 (13.5)
Method 2-from objectives	9 (9.4)
Method 3-most serious	15 (15.6)

393 **Abbreviation:** NA = not applicable.

ACCEPTED MANUSCRIPT

394 **Table 3. Changes in Primary Outcomes**

Change Type	**# of systematic reviews with ≥1 change(s) (%)***
New Primary Outcome(s)	6 (5.9)
Exclusion of Primary Outcome(s)	6 (5.9)
Upgrade of Primary Outcome(s)	6 (5.9)
Downgrade of Primary Outcome(s)	22 (21.8)
Change in Primary Outcome Definition	6 (5.9)
Change in Primary Outcome Measure	1 (1.0)
No Discrepancies	65 (67.7)

395 **Note:** *Does not add up to 100% because some systematic reviews included more than 1 primary
396 outcome.

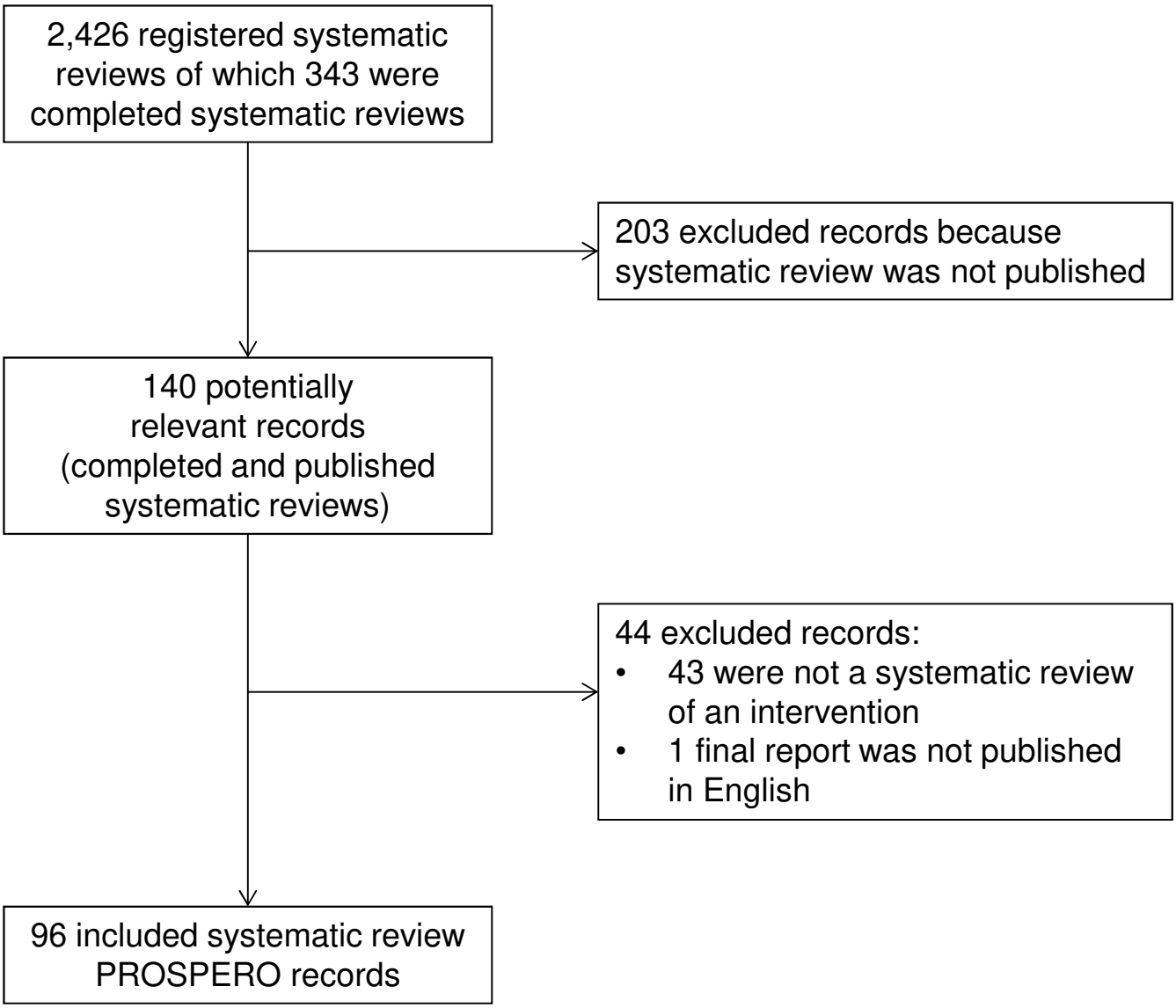
ACCEPTED MANUSCRIPT

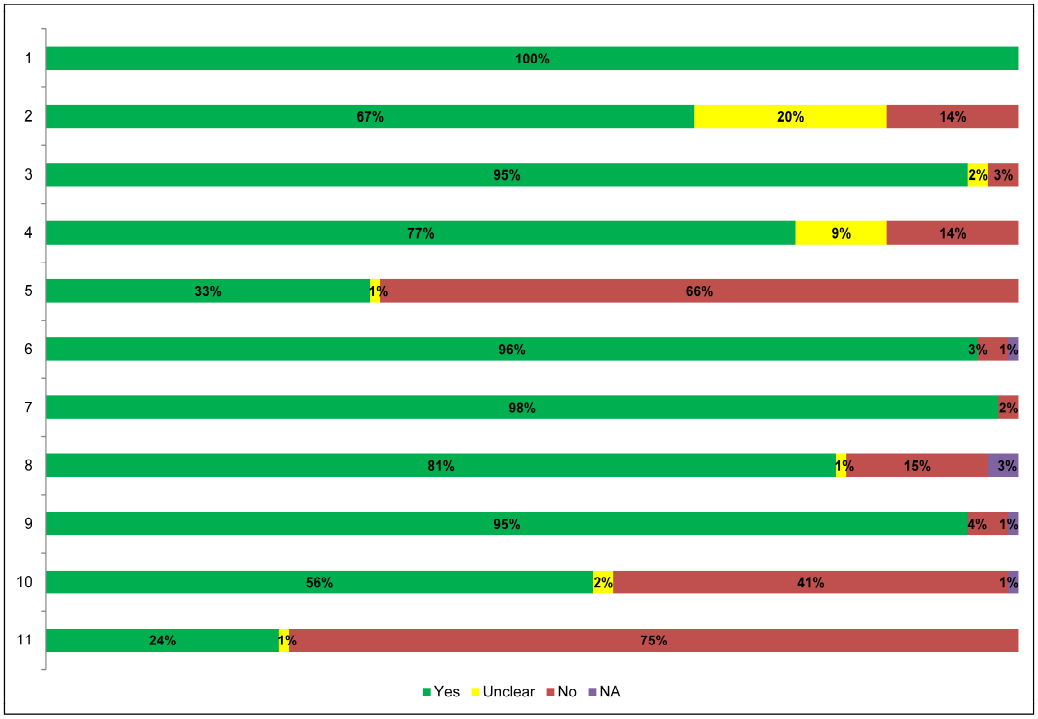
REFERENCES

1. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. In: Higgins J, Green S, eds: The Cochrane Collaboration; 2011.
2. Standards for Systematic Reviews. Institute of Medicine of the National Academies. : National Academy of Sciences; 2011.
3. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ. 2009;339:b2700.
4. Dwan K, Kirkham JJ, Williamson PR, Gamble C**.** Selective reporting of outcomes in randomised controlled trials in systematic reviews of cystic fibrosis. BMJ Open. 2013;3(6).
5. Kirkham JJ, Altman DG, Williamson PR**.** Bias due to changes in specified outcomes during the systematic review process. PLoS ONE. 2010;5(3):e9810.
6. Silagy CA, Middleton P, Hopewell S**.** Publishing protocols of systematic reviews: comparing what was done to what was planned. JAMA. 2002;287(21):2831-4.
7. Parmelli E., Liberati A., D'Amico R. Reporting of outcomes in systematic reviews: comparison of protocols and published systematic reviews. 15th Cochrane Colloquium. Sao Paulo; 2007.
8. Tricco AC, Tetzlaff J, Sampson M, Fergusson D, Cogo E, Horsley T, et al. Few systematic reviews exist documenting the extent of bias: a systematic review. J Clin Epidemiol. 2008;61(5):422-34.
9. Page MJ, McKenzie JE, Kirkham J, Dwan K, Kramer S, Green S, 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.
10. PROSPERO. York, England: Centre for Reviews and Dissemination, Univeristy of York.
11. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG**.** Epidemiology and reporting characteristics of systematic reviews. PLoS Med. 2007;4(3):e78.
12. Moher D, Dulberg CS, Wells GA**.** Statistical power, sample size, and their reporting in randomized controlled trials. JAMA. 1994;272(2):122-4.
13. Tricco AC, Tetzlaff J, Pham B, Brehaut J, Moher D**.** Non-Cochrane vs. Cochrane reviews were twice as likely to have positive conclusion statements: cross-sectional study. J Clin Epidemiol. 2009;62(4):380-6.e1.
14. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol. 2009;62(10):1013-20.
15. Mesgarpour B, Griebler U, Glechner A, Kien C, Strobelberger M, Van Noord MG, et al. Extended-release opioids in the management of cancer pain: a systematic review of efficacy and safety. Eur J Pain. 2014;18(5):605-16.
16. 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(9):e1000326.
17. Page MJ, McKenzie JE, Green SE, Forbes AB**.** An empirical investigation of the potential impact of selective inclusion of results in systematic reviews of interventions: study protocol. Syst Rev. 2013;2:21.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT




1
100%
2
67%
20%
14%
3
95%
2%
3%
4
77%
9%
14%
5
33%
1%
66%
6
96%
3%
1%
7
98%
2%
8
81%
1%
15%
3%
9
95%
4%
1%
10
56%
2%
41%
1%
11
24%
1%
75%
Yes
Unclear
No
NA

ACCEPTED