Author’s response to reviews

Title: Recommendations for Reporting of Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy: A Systematic Review

Authors:

Trevor McGrath (tmcgr043@uottawa.ca)
Mostafa Alabousi (malab101@uottawa.ca)
Becky Skidmore (bskidmore@rogers.com)
Daniel Korevaar (d.a.korevaar@amc.uva.nl)
Patrick Bossuyt (p.m.bossuyt@amc.uva.nl)
David Moher (dmoher@ohri.ca)
Brett Thombs (brett.thombs@mcgill.ca)
Matthew McInnes (mcinnes.matt@gmail.com)

Version: 1 Date: 21 Sep 2017

Author’s response to reviews:

September 15, 2017

Dear Dr. Page,

We are pleased to submit our revised manuscript: “Recommendations for Reporting of Systematic Reviews and Meta-Analyses of Diagnostic Test Accuracy: A Systematic Review”. We feel that the reviewer comments have been addressed and it is now a superior paper. Detailed response is outlined below.

Thank you on behalf of the authors,

Matthew McInnes MD FRCPC
Associate Professor of Radiology
University of Ottawa

Reviewer #1 Comments
1) The list of items pertains (correctly) to reporting of DTA reviews, but it is described elsewhere that the systematic review was of "existing guidance on quality of reporting and methodology" (Abstract) or "published articles pertaining to methodology or reporting quality" (Database search). The latter (guidance on DTA methodology) seems to me much wider than guidance on reporting. I would expect to find far more than 203 papers giving "guidance" on DTA review methodology, e.g. suggesting specific meta-analysis methodology in particular. Can the authors be much clearer about their inclusion criteria?

Reply: Thank you for your careful attention to detail. This wording has been clarified. The word “methodology” has been removed from “guidance on reporting and methodology” to reflect the nature of the items on the list.

2) There does not appear to be a list anywhere of the full 203 "included" papers (e.g. Figure 1).

Reply: Such a list has been added as Appendix 3.

3) The authors have undertaken a complex search strategy and detailed screening, which must have been a huge amount of work. However, in the end we find that "Items [in the final list] were taken from 19 unique sources". It is unclear how the 203 included papers were reduced to the final 19 sources.

Reply: The following text was added to the manuscript "Many of the 203 included results contained redundant information; one source was cited per item."

4) I was surprised that the list didn't include anything about reporting how multiple points on the ROC curve from an individual study were handled? (i.e. What was done if a study reported sensitivity and specificity at 2 or more different cut points?)

Reply: This is an excellent point and was not included on this pre-Delphi process list. Fortunately, this issue will be included as an item on the forthcoming final version of the PRISMA-DTA checklist.

Reviewer #2 Comments

1) "Methods - Database search": It is worth mentioning to readers that the Cochrane Methodology Register does not include any records published after July 2012 (see http://cmr.cochrane.org/).

Reply: We also searched the Cochrane Methodology Register in the Cochrane Library, which contains records published July 2012 and earlier, (Wiley version) on the same date.

The bolded text from the above sentence was added to the manuscript.
2) "Methods - Inclusion/Exclusion Criteria, Study Selection, and Data Extraction": It was a little unclear to me how the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy was used to generate potential items for PRISMA-DTA. That is, was the whole handbook read (cover-to-cover) by two authors independently? Also, the Handbook is incomplete currently, so it is worth mentioning this.

Reply: It was clarified that the Cochrane DTA Handbook was assessed in duplicate. It was noted that only the completed chapters of the Cochrane DTA Handbook were assessed.

3) "Methods - Inclusion/Exclusion Criteria, Study Selection, and Data Extraction": Are there any guidance documents that were excluded from the review? For example, the AHRQ Methods Guide for Comparative Effectiveness Research, the Institute of Medicine's 2011 Standards for Systematic Reviews, and the Centre for Reviews and Dissemination guidance, are not listed as sources. It would be good to explain why these were excluded.

Reply: Despite our comprehensive search strategy, they were not identified in the initial search; this may be related to the fact that they (2 of the 3) are targeted at intervention research/systematic reviews rather than those of diagnostic accuracy. The CRD document does have a chapter on diagnostic test reviews; however all of this guidance is included in other sources (e.g. Cochrane DTA handbook). Post-hoc evaluation of these sources does not identify any additional items that have not already been included.

The following text has been added to the manuscript:

Post-hoc assessment of the following items not included in the initial search was done: the Agency for Healthcare Research and Quality (AHRQ) Methods Guide for Comparative Effectiveness Research, the Institute of Medicine's 2011 Standards for Systematic Reviews, and the Centre for Reviews and Dissemination guidance [32-34]. No additional items were generated from these sources.

4) "Methods - Inclusion/Exclusion Criteria, Study Selection, and Data Extraction": The MECIR Standards were updated in July 2016 (see http://community.cochrane.org/mecir-manual). Suggest that you specify that the 2013 version was consulted for your review.

Reply: This is an excellent point. The citation we have used for MECIR indicates that the 2013 version was consulted.

5) "Methods - Inclusion/Exclusion Criteria, Study Selection, and Data Extraction": The methods used to generate a list of potentially relevant items are somewhat unclear. Were items in existing reporting guidelines extracted verbatim? Were items then revised so that they are more applicable to systematic reviews of DTA? How were items generated from the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy (given this is a handbook, not a
structured list of items)? How were items from tools for assessing risk of bias or quality in systematic reviews generated (e.g. AMSTAR, ROBIS), given these include a list of questions for users of systematic reviews to answer. And how were items in reporting guidelines for primary DTA studies handled (e.g. STARD 2015, STARD for Abstracts)?

Reply: The following was added to the section: “Items deemed relevant may have had wording changed from the original source to make them more applicable to systematic reviews of diagnostic test accuracy and/or broken into multiple sub-items to facilitate the Delphi process for PRISMA-DTA.”

6) "Results - Identification of Potentially Relevant Items": Some more detail on how items were divided into categories could be provided. For example, was categorisation done by one or two authors? Were categories pre-specified? Are categories aligned with the existing PRISMA Statement?

Reply: After searching the existing literature and guidance documents, a preliminary list of 64 unique items was compiled and divided into the following categories mirroring the PRISMA statement:

The bolded text from the above sentence was added to the manuscript.

7) "Results - Identification of Potentially Relevant Items": The paragraph starting, "Items were taken from 19 unique sources…", is a little confusing. By this I assume you mean that these 19 sources are different to the 203 articles identified from the search of bibliographic databases? If so, then what was obtained from the 203 articles? Also, were the "eight research articles [6, 17, 34-39], two reviews [2, 40], two DTA statistical methodology overviews [41, 42], and one conference abstract [43]" identified from the search of bibliographic databases (that is, included in the list of 203 articles)?

Reply: Items were taken from 19 unique sources with publication dates between 2007 and 2016; a combination of guidance documents and some of the 203 search results.

The bolded text from the above sentence was added to the manuscript

8) "Discussion": The discussion could be extended. Currently there is little comparison with existing research or similar studies of this nature, and not much consideration of the implications of the research.

Reply: Discussion has been expanded. Thank you.