Author’s response to reviews

Title: Risk of Bias Tools in Systematic Reviews of Health Interventions: An Analysis of PROSPERO-Registered Protocols

Authors:

Kelly Farrah (kfarr081@uottawa.ca)
Kelsey Young (kelsey.young@canada.ca)
Matthew Tunis (matthew.tunis@canada.ca)
Linlu Zhao (linlu.zhao@canada.ca)

Version: 1 Date: 12 Jun 2019

Author’s response to reviews:

The authors would like to thank the reviewers for their time and very helpful comments. Based on these comments, we have made major revisions to our manuscript and have provided responses explaining our rationale below. Revisions are highlighted in the manuscript using the “track changes” function in Word. A clean copy of the manuscript has also been submitted. Thank you for your further consideration of our manuscript.

REVIEWER 1 COMMENT 1:

Page 7, Line 142: As you haven’t cross checked, it is a limitation. May be the authors just named the tool in their review and did not use it. Is it possible to check randomly (around 10%)?

AUTHOR RESPONSE:

We did not further verify the text of the protocol records in the longitudinal analysis, as the intent of this analysis was to estimate the trends in usage of selected risk of bias tools since the inception of PROSPERO. Therefore, we felt that using raw search returns from the use of risk of bias terms act as an adequate proxy for the trend of the planned use of a tool over time in the PROSPERO database.

Regarding the issue of concurrence between a tool named in the PROSPERO protocol and its actual usage in the published review, this is an interesting topic, but we feel that it is outside the scope of the current analysis and a unique topic to be tackled on its own. Furthermore, as we conducted the cross-sectional analysis of protocols from 2018, we feel that it is a reasonable assumption that insufficient time has elapsed to account for the time to conduct, peer review, and publish a large majority of these reviews, rendering a concurrence analysis at this point less
meaningful and would risk the findings becoming dated if we waited for sufficient elapsed time for publication of these reviews.

REVIEWER 1 COMMENT 2:

Page 9, Line 183, Table 2: Is there any difference between "Cochrane ROB tool" and "Cochrane ROB 2 tool"? Same question for "Cochrane ROB tool" and "Cochrane handbook". Should you merge these during analysis?

AUTHOR RESPONSE:

Yes, there is a difference between Cochrane RoB tool and Cochrane RoB 2 Tool. The latter is an updated version of the Cochrane RoB Tool and was first published in 2016. We separated these two versions as it gives readers a sense of the current (2018) state of uptake of the version 2.0 of the Cochrane RoB tool, with the caveat that the frequency for version 2.0 of the tool may be underrepresented, as authors who planned to use the Cochrane RoB tool may actually mean version 2.0 of the tool, but didn’t specify the version in their PROSPERO protocol. This explanation is now included in the first paragraph of the Discussion section (Page 14, Lines 302-306 in the tracked changes document).

The Cochrane Handbook is not a risk of bias tool, but rather it presents an approach to evidence synthesis for Cochrane reviews. This and the GRADE approach were identified by some PROSPERO protocol authors as “tools” for assessing risk of bias in included interventional studies. We felt that these should be included to illustrate that some authors may be misidentifying an approach as a tool, or may not be sufficiently specific in their protocols by identifying an evidence synthesis approach which may include suggested tools. We clarified this point (i.e., approach and not risk of bias tool) in the footnote to Table 2.

REVIEWER 1 COMMENT 3:

Page 14, Line 301: Cochrane handbook has described tools for non-randomized trials and they have mentioned Down and Black and NOS as most useful tool based on a review.

AUTHOR RESPONSE:

We agree that there is guidance for the quality assessment of non-randomized studies, like in the Cochrane Handbook. We clarified the statement to say that some protocols may have inappropriately applied the Cochrane RoB Tool intended for RCTs for the quality appraisal of non-randomized studies, and this suggests an opportunity for more education and awareness around the appropriate use of study design-specific quality appraisal tools.
REVIEWER 1 COMMENT 4:
Too much acronym sometimes distracted the reading

AUTHOR RESPONSE:
Thank you for this observation. We agree that excessive acronym usage distracted from reading and have removed the use of the following acronyms in the text: “SR” for systematic review, “NOS” for Newcastle-Ottawa Scale, “HTA” for health technology assessments, and “RoB” for risk of bias only in the context of “RoB tools”.

REVIEWER 2 COMMENT 1:
The authors should provide a definition of risk of bias in the introduction.

AUTHOR RESPONSE:
Cochrane Collaboration’s definition for risk of bias is now included in the Introduction section (Page 3, Lines 56-57 in the tracked changes document).

REVIEWER 2 COMMENT 2:
The authors refer to their study as a retrospective review of the Prospero registry for ROB of NRS in the title, and this should be explicitly stated in the last section of the introduction.

AUTHOR RESPONSE:
Because our analysis includes a cross-sectional look at the most commonly cited RoB tools in 2018 and a retrospective look at the trend in use of these commonly cited tools over the past 8 years, we revised the title of the article to be more reflective of this broader analysis. The last paragraph of the Introduction section containing objectives of the study was also revised accordingly for consistency and clarification. The revised title is:

Risk of Bias Tools in Systematic Reviews of Health Interventions: An Analysis of PROSPERO-Registered Protocols

REVIEWER 2 COMMENT 3:
Can the authors include a list of terms they searched for the 2018 records in Prospero.

AUTHOR RESPONSE:
PROSPERO-defined search filters were applied to retrieved potentially relevant records from 2018 in PROSPERO. Therefore, no search terms per se were used to conduct the search. The Methods section has been revised to more explicitly state that database filters were used to conduct the search and that the exact search strategy used was:
( Intervention): RT NOT Cochrane: DB WHERE CD FROM 01/01/2018 TO 12/10/2018).

REVIEWER 2 COMMENT 4:

Can the authors comment on how records were screened. Was this completed in duplicate, how were disagreements resolved?

AUTHOR RESPONSE:

This detail has been added to the “Data Source and Sample Selection” section of Methods. One reviewer screened the records for inclusion as we felt that the risk of error in screening and data extraction was minimal due to the nature of the data source (e.g., consistent structure across protocols, discrete identifiable tools).

REVIEWER 2 COMMENT 5:

It seems strange the authors would not search all 4000 protocols, as the information they are abstracting is relatively little. Can the author provide a stronger justification for a 10% sample? Why not 50% or 25% since they are ascertaining trends?

AUTHOR RESPONSE:

As our analysis examines commonly cited RoB tools in contemporary use, we conducted a cross-sectional analysis of protocols in 2018. Therefore, to achieve our aim of identifying common tools in this time period (and not undertake a census of all tools that are planned to be used) we felt that a 10% random sample would adequately capture the common tools cited for use. The rationale for choosing this 10% random sample has been strengthened in the Methods section.

REVIEWER 2 COMMENT 6:

Lines 124-125 is confusing. Can the authors clarify what is meant by "produced by the same organization" or provide a specific example.

AUTHOR RESPONSE:

Here we refer to sets of related checklists assessing risk of bias for different study designs created by the same authors/producing body, for example Joanna Briggs Institute (JBI) checklists. A systematic review protocol may have simply said that they were using “JBI checklists” to evaluate risk of bias for all study designs included. In such cases, we were unable to determine the specific number of risk of bias checklists used in the review.

REVIEWER 2 COMMENT 7:

Lines 125-127 are redundant as this was specified in the previous section on sample selection.
AUTHOR RESPONSE:

We agree with deletion of the redundant text that re-iterates the selection criteria.

REVIEWER 2 COMMENT 8:

Longitudinal analysis - It does not seem methodological to search only 10% of the records and then only search for those tools retrospectively. I think this does not truly highlight actual trends in ROB for NRS over time can the authors address this or justify in the discussion section. Lines 139-141 are redundant and specified in the previous section on sample selection.

AUTHOR RESPONSE:

Our intent with the longitudinal analysis was to look at the trends in use of the most commonly cited tools identified in the 2018 cross-sectional analysis. As there are many RoB tools in existence, we wanted to use a more objective method to curate a list of tools and look at trends in their use. We felt that focusing on the more common tools in contemporary use is more relevant to the evidence synthesis community at present time to facilitate discussions around appropriate use of RoB tools for NRS in particular.

REVIEWER 2 COMMENT 9:

Did the authors contact authors for the 4 protocols that were unclear?

AUTHOR RESPONSE:

We made the decision to not contact the authors for the 4 protocols that were unclear on their RoB methodology, as we felt that identifying the frequency of protocols with unclear tools in the cross-sectional analysis provides a picture of how prevalent a lack of clearly defined review protocol is in PROSPERO. Furthermore, we felt that clarity on these few unclear protocols would not substantially impact the findings on the most common tools cited for use for both RCT and NRS.

REVIEWER 2 COMMENT 10:

For the studies reporting on RCTs does this include quasi-RCT as well, or are they just RCTs?

AUTHOR RESPONSE:

We grouped the protocols looking at RCTs or quasi-RCTs together, as the primary distinction of interest in our analysis is RCT (including quasi-RCT) and NRS.

REVIEWER 2 COMMENT 11:

I am unclear why the Cochrane handbook is listed as a tool in Table 2. The handbook is not an ROB tool. Similar concerns about the AND/ADA Quality Checklist or GRADE.
AUTHOR RESPONSE:

We agree that these are not RoB tools, but were identified as such by the protocol authors. We felt that there was value in presenting these cited “tools” as described by the protocol authors in order to point out that some protocol authors are not being specific in identifying RoB tools or are inappropriately applying RoB methods. We have added text explaining this in a footnote to Table 2.

REVIEWER 2 COMMENT 12:

Similar concerns for table 3 - Strobe statement is not an ROB tool or the CASP checklist. Why are these included in the results, as authors stated reporting guidelines and checklists were excluded from results?

AUTHOR RESPONSE:

Echoing our response to the previous comment, we agree that the STROBE statement is a reporting guideline, not an RoB tool, but was identified by protocol authors as such. We have added text explaining this in a footnote to Table 3. We did consider the set of CASP checklists as potential tools for RoB assessment in the longitudinal analysis, although that they are not validated tools and were originally designed for educational purposes, they are referred to as “critical appraisal tools.”

REVIEWER 2 COMMENT 13:

The authors then state in the 'Annual trends' section they removed Cochrane and GRADE. For consistency please also remove from ROB tools identified in 2018.

AUTHOR RESPONSE:

For the protocol author-identified “tools” such as GRADE, CASP, STROBE etc. that are not actually RoB tools, we were interested in presenting these in the cross-sectional analysis to show that some protocol authors may be misusing tools to appraise RoB or were not being specific in the RoB to be used by describing an approach rather than naming a tool. However, for the longitudinal analysis, we excluded these “tools” from analysis as we are interested in looking at the trend in use of protocol-cited tools that are in current common usage, and inclusion of the much less frequently mis-used/identified tools would not be informative in this context.

REVIEWER 2 COMMENT 14:

For lines 223-237 - please specify if the RCTs using Cochrane ROB included quasi-studies. If they are RCTs only, I am not sure why they would be included, as the focus is on NRS.

AUTHOR RESPONSE:
We would like to clarify that our analysis was not restricted to NRS designs, but also included RCTs. Our discussion focused more on NRS as the findings for NRS was much more heterogeneous compared to RCT. For analytical purposes, true RCTs and quasi-RCTs are grouped together compared to NRS and this is noted in the Data Extraction section of Methods. The manuscript has been revised in various relevant sections to provide more clarity that the analysis looks at planned use of RoB tools for both RCTs and NRS.

REVIEWER 2 COMMENT 15:

The authors should include a limitations section in their discussion, which should address the 10% sample selection and not a true indication of shifts in trends used by authors over the last 8 years of the Prospero database.

AUTHOR RESPONSE:

We added the limitation of the trend analysis in the Discussion section (Page 16, Lines 348-351 in the tracked changes document).

REVIEWER 2 COMMENT 16:

Why did the authors not consider looking at published SR protocols? Published protocols are more likely to provide additional around ROB tools. Can the authors please comment on this?

AUTHOR RESPONSE:

We decided to focus on PROSPERO protocols as not all review protocols will be published in journals; in other words, although both PROSPERO and published protocols are subsets of all published reviews, PROSPERO protocols may be more representative compared to published protocols. Furthermore, PROSPERO protocols may be more recent compared to protocols published in a journal and using the former aids in our objective of finding the most recent usage information. Another reason we focused on PROSPERO protocols is that the PROSPERO registry is designed such that we can specifically search for tool names within the “Risk of bias (quality) assessment” field of the protocols; therefore, this is a much more efficient and sensitive approach than searching traditional literature databases for published protocols. Finally, if a review protocol is leveraging an existing tool, such as the Cochrane RoB Tool, they would tend to not explain the tool in more detail but rather cite out for additional details, and therefore, unless protocol authors used a custom tool, the information on an existing tool contained in a PROSPERO protocol should provide sufficient information for this analysis compared to a published protocol in most cases.
The intent of our study was not to evaluate consistency between tools identified for use in PROSPERO protocols versus the tools actually used in final publication. This would be an excellent topic for a future research study.

REVIEWER 2 COMMENT 17:

What about the validity and reliability of the most frequently used tools identified in the retrospective search. Can the authors include this information?

AUTHOR RESPONSE:

The most frequently used tools for NRS such as NOS, ROBINS-I and MINORS have not been appropriately validated. We added this information in the Discussion section (Page 15, Lines 323-325 in the tracked changes document).