PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	The role of pragmatism in explaining heterogeneity in meta-analyses of randomized trials: a protocol for a cross-sectional methodological review
AUTHORS	Aves, Theresa; Allan, Katherine; Lawson, Daeria; Nieuwlaat, Robby; Beyene, Joseph, Mbuagbaw, Lawrence.

VERSION 1 - REVIEW

REVIEWER	G.J. Melendez-Torres Warwick Medical School, University of Warwick, UK
REVIEW RETURNED	14-Jun-2017

GENERAL COMMENTS	This is, in the main, a strong protocol for an interesting study. I believe some details need more thought, and I have suggested several analytic possibilities for this study as well. --Can you discuss whether or not the PRECIS-2 tool has shown good psychometric properties? You hint at this but it could do with clearer surfacing. This may also be a good secondary objective for your study if you conclude the tool could do with more validating. --It seems that your focus on pragmatism to explain heterogeneity is a little cart-before-horse. How do you plan on assessing the utility of the tool in critical appraisal and reviews before using it in meta-regressions? It would be good to plan on offering structured insights in this respect as well. --Objective 3 does not feel clearly anticipated by the introduction. Could you offer more context for this? --Can you please provide a specific search string as you will run it in the Cochrane Library? I wonder if you even need a search string at all--just search for non-empty reviews of interventions. That seems like it would not involve a great deal more work than setting up a search and then running and debugging it. --I would suggest looking at individual PRECIS domains as meta-regressors as well. --While I understand the rationale behind choosing reviews with >/=10 studies in the primary meta-analysis, I would add the caveat that this relates to >/=10 studies considered in one pooled effect relating to the primary outcome. Reviews may subgroup within primary outcome based on time to follow-up.

	--You might also consider looking at change in I2 to quantify the explanatory value of PRECIS.

REVIEWER	Dr. Tolulope Sajobi University of Calgary, Calgary, Canada
REVIEW RETURNED	23-Jun-2017

GENERAL COMMENTS	This study provides the protocol of a meta-analysis that aims to investigate the role of pragmatism as a source of heterogeneity in systematic reviews by (1) identifying systematic reviews with meta-analyses of RCTs that have moderate to high heterogeneity; (2) determining how much of this heterogeneity is explained by the amount of pragmatism; (3) and exploring the risk of bias assessments in pragmatic trials. The results of this study will have important implications for the synthesis of evidence from randomized controlled trials and observational studies. Nevertheless, I have few concerns about this protocol in its current form. In the abstract, the authors state "…....based on the amount of pragmatism…". Please change the word "amount" to "degree" I find the introduction section to be rather too brief and the study rationale not elaborated on. The authors need to provide a detail review of other instruments used to assess pragmatism of trials. For example, instruments such as pragmascope (Tosh et al) should be described. The introduction lacks details on systematic reviews that has used PRECIS tool in their review. While the authors are right to suggest that this is not the first study to assess the using of PRECIS tools for aiding systematic reviews, they need to describe these literature in detail. Please see the following papers for more details. Koppenaal et al (2011), Yoong, Wolfenden, Clinton-McHarg, et al (2014) Random selection of 10 systematic reviews might introduce variations in the relative contribution of pragmatism on heterogeneity among the studies in a systematic review. The authors should probably consider selecting systematic reviews across a similar theme. For example, systematic reviews on cardiovascular disease, or systematic reviews on children, etc. The authors plan to assess the relative contribution of pragmatism on the amount of heterogeneity in treatment effect among the studies. But stopped short of discussing how they plan pool evidence about the meta-analysis of relative contribution of pragmatism in each systematic review across systematic reviews. Please provide more details. In the discussion section, please discuss the potential limitations of this study as identified in the "strengths and limitation" section of the manuscript

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:
Can you discuss whether or not the PRECIS-2 tool has shown good psychometric properties? You hint at this but it could do with clearer surfacing. This may also be a good secondary objective for your study if you conclude the tool could do with more validating.

Response: Following the submission of our protocol on May 21, 2017, a publication focusing on inter-rater reliability and discriminate validity of PRECIS-2 was published online in the Journal of Clinical Epidemiology by Loudon K, et al. (article currently in press). In light of these new results, we have incorporated the main findings in the introduction of our protocol to provide further context. Please refer to page 5 for the revisions.

In summary, 19 experienced trialists and methodologists agreed to review 10-15 trial protocols and assess them according to the 9 PRECIS-2 domains. Inter-rater reliability ranged from 0.24 (flexibility (adherence)) to 0.94 (eligibility) with 7 of 9 domains having an intraclass correlation coefficient >0.65 suggesting substantial agreement according to Landis and Koch, 1977. Discriminate validity was assessed using the area under the curve for each PRECIS-2 domain which ranged from 0.57 (organization) to 0.75 (primary outcome) suggesting fair discriminability in 4 of the 9 domains (≥0.7 to 0.8) according to Kleinbaum and Klein, 2010. Discriminate validity was lowest in the setting and organization domains of PRECIS-2, both with an area under the curve <0.6.

Additionally, we will include the assessment of inter-rater reliability as an additional objective in our research. Loudon et al. used study protocols to assess inter-rater reliability, however we will be using primary studies of systematic reviews and will assess inter-rater reliability for each PRECIS-2 domain as well as the overall summary score. As study protocols are not always available in a published format, primary studies may be more applicable to assess PRECIS-2 and will provide reliability information in a systematic review setting. Please see the abstract and objectives sections (pages 2 and 5, respectively) for the inclusion of this objective.

It seems that your focus on pragmatism to explain heterogeneity is a little cart-before-horse. How do you plan on assessing the utility of the tool in critical appraisal and reviews before using it in meta-regressions? It would be good to plan on offering structured insights in this respect as well.

Response: Thank you for your comment. We have provided further detailed information of how we intend to assess the PRECIS-2 tool which includes an evaluation of inter-rater reliability (which will serve as a surrogate measure of agreement of our data extraction), a categorical schema for subgroup analyses, and meta-regression. Previous literature has focused on the utility of PRECIS in systematic reviews, which has been regarded as useful and important. Our goal is to take the research one step further to determine how it may be a potential source of heterogeneity in a systematic review setting.

Objective 3 does not feel clearly anticipated by the introduction. Could you offer more context for this?

Response: Thank you for your comment. We agree and have removed this objective from the protocol as it lends itself to an entirely separate research project which we will conduct at a later time.

Can you please provide a specific search string as you will run it in the Cochrane Library? I wonder if you even need a search string at all--just search for non-empty reviews of interventions.

That seems like it would not involve a great deal more work than setting up a search and then running and debugging it.

Response: The search string is as follows: randomize:ti,ab,kw or RCT:ti,ab,kw; Online Publication Date from Jan 2014 to Jan 2017 (Word variations have been searched). This search yielded 2617 citations which were imported into Distiller SR where they were screened for inclusion criteria by 2 independent reviewers (TA, KA). We have clarified the search strategy in the protocol, please see the methods section (page 6) for the revision.

I would suggest looking at individual PRECIS domains as meta-regressors as well.

Response: Thank you for your comment, we will consider looking at individual PRECIS-2 domains as meta-regressors. We have included this in the data analysis section (page 7).

While I understand the rationale behind choosing reviews with >/=10 studies in the primary meta-analysis, I would add the caveat that this relates to >/=10 studies considered in one pooled effect relating to the primary outcome. Reviews may subgroup within primary outcome based on time to follow-up.

Response: We agree with your comment and have revised the methods to incorporate this caveat. The methods now reads, “Inclusion criteria will include systematic reviews of RCTs from any Cochrane Review Group with at least 10 studies considered in one pooled effect relating to the primary outcome and moderate to substantial heterogeneity (I2≥ 50%)."

We do concur that reviews may subgroup based on time to follow-up or other a priori specified measures that may explain heterogeneity. We will discuss how PRECIS-2 may lend itself to subgroup considerations when heterogeneity has not been explained by subgroups as specified a priori by the Cochrane review authors.

You might also consider looking at change in I2 to quantify the explanatory value of PRECIS.

Response: Thank you for your comment, we agree and will consider looking at change in I2 to quantify the explanatory value of PRECIS-2.

Reviewer 2:
In the abstract, the authors state “…....based on the amount of pragmatism…”. Please change the word “amount” to “degree”

Response: We have changed “amount” to “degree” in the abstract and all sections in the body of the text where we referred to the “amount of pragmatism”.

I find the introduction section to be rather too brief and the study rationale not elaborated on. The authors need to provide a detail review of other instruments used to assess pragmatism of trials. For example, instruments such as pragmascope (Tosh et al) should be described.

Response: Thank you and we agree with your comment. We have revised the introduction section of the protocol and have discussed additional instruments such as the Pragmascope. Additionally, we have provided further study rationale based on your feedback. Please refer to the revised introduction section (page 3) for these changes.

The introduction lacks details on systematic reviews that has used PRECIS tool in their review. While the authors are right to suggest that this is not the first study to assess the using of PRECIS tools for aiding systematic reviews, they need to describe these literature in detail. Please see the following

papers for more details. Koppenaal et al (2011), Yoong, Wolfenden, Clinton-McHarg, et al (2014)

Response: We agree with your comment and have included your suggested citations, as well as an additional citation from Witt et al. 2012. While we are certainly not the first to assess PRECIS in a systematic review setting, we are the first to assess PRECIS-2 using their formalized scoring system from 1 to 5. Koppenaal et al., Yoong et al., and Witt et al. applied the original 10 domains of PRECIS to studies using rating systems of 0-4 or 1-5 with the lowest number representing more explanatory trials and the highest number representing more pragmatic trials. Please see the revised introduction section (pages 3-5) for the inclusion of the references.

Random selection of 10 systematic reviews might introduce variations in the relative contribution of pragmatism on heterogeneity among the studies in a systematic review. The authors should probably consider selecting systematic reviews across a similar theme. For example, systematic reviews on cardiovascular disease, or systematic reviews on children, etc.

Response: The rationale behind a random selection of 10 systematic reviews was for two reasons, 1. to reduce bias in the systematic review selection process and 2. we believe that PRECIS-2 may be applicable across a wide range of disease conditions and types of interventions. We did start off by limiting our search of systematic reviews to interventions in cardiovascular disease, however our preliminary search did not yield enough systematic reviews meeting inclusion criteria for us to proceed this way. We re-evaluated and decided that limiting ourselves to systematic reviews of a similar theme may not allow us to fully explore the potential applicability of the tool. Therefore, we decided to include systematic reviews from any review group in the Cochrane Database. We do recognize that there may be differences in the contribution of pragmatism on heterogeneity among primary studies within a systematic review. We are prepared to discuss these potential differences once we have analyzed the results.

The authors plan to assess the relative contribution of pragmatism on the amount of heterogeneity in treatment effect among the studies. But stopped short of discussing how they plan pool evidence about the meta-analysis of relative contribution of pragmatism in each systematic review across systematic reviews. Please provide more details.

Response: Thank you for your comment. In the first instance, we are going to investigate the effects of pragmatism within each systematic review. Secondly, we will explore this relationship across systematic reviews with similar outcome types (i.e. continuous, binary or time-to-event), using the systematic review as a grouping factor. This has been clarified in the text, please see the data analysis section (page 7) for this revision.

In the discussion section, please discuss the potential limitations of this study as identified in the "strengths and limitation" section of the manuscript

Response: We have included the limitations of the study as outlined in the strengths and limitations section of the paper. Please see the discussion and dissemination section (page 8) for these changes.

Once again, we thank you for your review of our research protocol. We hope that you agree with our revisions and find our responses satisfactory.

VERSION 2 – REVIEW

REVIEWER	G.J. Melendez-Torres Warwick Medical School, University of Warwick, UK
REVIEW RETURNED	28-Jul-2017

GENERAL COMMENTS	Authors have satisfactorily addressed all comments.

REVIEWER	Dr. Tolulope Sajobi University of Calgary, Canada
REVIEW RETURNED	21-Jul-2017

GENERAL COMMENTS	The authors have adequately addressed the concerns raised