OPEN ACCESS

Knowledge user survey and Delphi process to inform development of a new risk of bias tool to assess systematic reviews with network meta-analysis (RoB NMA tool)

Carole Lunny [1,2] Areti Angeliki Veroniki,[1,3]
Brian Hutton [4,5] Ian White,[6] JPT Higgins,[7] James M Wright,[2]
Ji Yoon Kim,[8] Sai Surabi Thirugnanasampanthar,[9]
Shazia Siddiqui,[1] Jennifer Watt [10] Lorenzo Moja,[11]
Nichole Taske,[12] Robert C Lorenz [13] Savannah Gerrish,[14]
Sharon Straus,[1,10] Virginia Minogue,[15] Franklin Hu,[16]
Kevin Lin,[16] Ayah Kapani,[16] Samin Nagi,[16] Lillian Chen,[16]
Mona Akbar-nejad,[16] Andrea C Tricco[1,3]

ABSTRACT

Background Network meta-analysis (NMA) is increasingly used in guideline development and other aspects of evidence-based decision-making. We aimed to develop a risk of bias (RoB) tool to assess NMAs (RoB NMA tool). An international steering committee recommended that the RoB NMA tool to be used in combination with the Risk of Bias in Systematic reviews (ROBIS) tool (i.e. because it was designed to assess biases only) or other similar quality appraisal tools (eg, A MeaSurement Tool to Assess systematic Reviews 2 [AMSTAR 2]) to assess quality of systematic reviews. The RoB NMA tool will assess NMA biases and limitations regarding how the analysis was planned, data were analysed and results were presented, including the way in which the evidence was assembled and interpreted.

Objectives Conduct (a) a Delphi process to determine expert opinion on an item's inclusion and (b) a knowledge user survey to widen its impact.

Design Cross-sectional survey and Delphi process.

Methods Delphi panellists were asked to rate whether items should be included. All agreed-upon item were included in a second round of the survey (defined as 70% agreement). We surveyed knowledge users' views and preferences about the importance, utility and willingness to use the RoB NMA tool to evaluate evidence in practice and in policymaking. We included 12 closed and 10 open-ended questions, and we followed a knowledge translation plan to disseminate the survey through social media and professional networks.

Results 22 items were entered into a Delphi survey of which 28 respondents completed round 1, and 22 completed round 2. Seven items did not reach consensus in round 2. A total of 298 knowledge users participated in the survey

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ The development of new tools to inform evidence-based medicine requires the feedback of knowledge users and experts.

⇒ The purpose of the knowledge user survey and Delphi process was to ask respondents about the structure of a proposed tool for assessing biases in an network meta-analysis (NMA), and about the concepts related to bias in NMAs to potentially include in the tool.

WHAT THIS STUDY ADDS

⇒ The majority of knowledge users and Delphi respondents preferred a tool to assess the bias in an individual NMA's results and in authors' conclusions.

⇒ Delphi respondents agreed to potentially include 15 out of 22 concepts about bias in NMAs.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This Delphi and knowledge user surveys inform the development of a new risk of bias NMA tool.

(14% respondent rate). 75% indicated that their organisation produced NMAs, and 78% showed high interest in the tool, especially if they had received adequate training (84%). Most knowledge users and Delphi panellists preferred a tool to assess both bias in individual NMA results and authors' conclusions. Response bias in our sample is a major limitation as knowledge users working in high-income countries were

more represented. One of the limitations of the Delphi process is that it depends on the purposive selection of experts and their availability, thus limiting the variability in perspectives and scientific disciplines.

Conclusions This Delphi process and knowledge user survey informs the development of the RoB NMA tool.

Introduction

Guidance about how to systematically develop quality and bias assessment tools is well established,[1] [2] multistaged and includes the involvement of knowledge users and experts. The benefits of engaging knowledge users in tool development is a key factor associated with knowledge translation and the reduction of research waste.[3–8] Specifically, the benefits include: greater public acceptance[9]; identifying and prioritising topics for research[10]; providing feedback on the tool's usability[10]; wider dissemination, uptake and communication of findings[10] and increased likelihood of impact.[10] [11] Identifying an external group of experts to obtain a multitude of perspectives will produce a more valid tool than a judgement given by an individual expert, or a group of experts heavily involved in the development process. Engaging with knowledge users and experts during development ensures that new tools will be relevant and applicable.

The risk of bias in network meta-analysis (RoB NMA) tool project aims to develop the first tool to assess risk of bias in a review with network meta-analyses (NMAs). We intended the RoB NMA tool to be used in combination with ROBIS[12] (which we recommend as it was designed to assess biases specifically) or other similar tools (eg, AMSTAR 2[13]) to assess quality of systematic reviews. The RoB NMA tool will assess NMA biases and limitations regarding how the analysis was planned, data were analysed and results were presented, including the way in which the evidence was assembled and interpreted. Our proposed RoB NMA tool has several uses. It can help knowledge users: (i) decide whether to believe the results from a single NMA and (ii) help choose between NMAs based on their risk of bias.

Development of the tool follows five stages. In the first stage, we conducted and published a methodological review to identify items related to bias in NMAs[14]; second, the steering group made conceptual decisions about the type of tool that will be developed, and refined the items into concepts from the methodological review; third, a knowledge user survey was developed to solicit feedback on the structure of the tool from potential users and fourth, expert opinion was obtained through a Delphi survey to select and define the concepts. The final and future stage will involve compiling the items into a tool; and conducting pilot testing to refine the items in the tool.

In this paper, we report on the knowledge user survey and Delphi process. We define 'knowledge user' as an individual who is likely to be able to use research results to make informed decisions about health policies, programmes and/or practices.[15] A knowledge user can be, but is not limited to, a practitioner, a policy maker, an educator, a decision maker, a healthcare administrator, a community leader or an individual in a health charity, patient group, private sector organisation or media outlet.[15] Our definition of experts is based on an individual's scientific/professional expertise, in our case in methods for NMAs, bias in systematic reviews with or without NMAs and risk of bias tool development.

The purpose of the survey was to ask knowledge users about the structure of our proposed tool and about their potential use of the tool in evidence-informed practice, policymaking, guideline development or research. We also aimed to conduct a multiround Delphi process to solicit expert opinion about concepts to potentially include in the tool.

Methods

Management of the project

At the start of our project to develop a risk of bias tool for NMAs, we first convened a steering group of nine experts in NMA, bias and tool development (online supplemental appendix A).[12] The steering group is responsible for the management of the project and has executive power over all decisions related to the proposed tool, which is still under development.

Protocol

We uploaded our study protocol on the Open Science Framework at https://osf.io/da4uy/. The knowledge user survey complied with the Checklist for Reporting Results of Internet E-Surveys (online supplemental appendix B).[16] Important definitions are found in box 1.

A cross-sectional survey design was used for the knowledge user survey using Qualtrics.[17] [18] Unique site visitors were identified via IP address and personal information was collected on a voluntary basis from respondents (no incentives were offered). Knowledge users and experts were identified using a purposive sampling strategy.

Knowledge user survey

Design

An English-language survey with 15 (12 closed and 10 open-ended) questions was developed by the investigative team (online supplemental appendix C). Five authors piloted the survey and modified it iteratively to improve content validity. Respondents were allowed to skip questions they did not wish to answer. The knowledge user survey ran from June 28 to 1 August 2021.

There were two parts to the survey: (1) demographic information and information about whether the knowledge users' organisation used or produced NMAs; (2) purpose of the RoB NMA tool, namely whether knowledge users preferred to assess the bias in the results, the authors' conclusions of an NMA or both. Further sections asked about interest and engagement in development, piloting, dissemination and training.

Email list development

We created an email list of journal editors publishing NMAs, using one bibliometric study of NMAs.[19] From this list, we extracted the journal names, and names of authors of NMAs. We also developed a list of organisations and institutions producing NMAs (online supplemental appendix D). We also included in the email list respondents from a UBC Methods Speaker Series on evidence synthesis methods (https://www.ti.ubc.ca/2022/01/01/methods-speaker-series-2022/).

Dissemination

All potential survey respondents were sent an email describing the purpose of the study, requesting their participation and providing a link to the survey (online supplemental appendix E). A knowledge translation plan was followed to disseminate and advertise the survey (online supplemental appendix F). We used twitter cards (ie, advertisements with pictures) and targeted hashtags to increase awareness of the survey (see the Twitter Campaign in online supplemental appendix G). In addition, we advertised through the e-newsletters of Knowledge Translation Canada, SPOR Evidence Alliance and Therapeutics Initiative.

Box 1   Important definitions

Network meta-analysis (NMA)

We adopted a broad definition of an NMA as a method that aims to, or intends to, synthesise simultaneously the evidence from multiple studies investigating more than two healthcare interventions of interest. We used the Cochrane Handbook definition of an NMA: '*Any set of studies that links three or more interventions via direct comparisons forms a network of interventions. In a network of interventions there can be multiple ways to make indirect comparisons between the interventions. These are comparisons that have not been made directly within studies, and they can be estimated using mathematical combinations of the direct intervention effect estimates available*'.[37] A network is composed by at least three nodes (interventions or comparators) and these are connected (graphically depicted as lines/edges) when at least one study compares the underlying two interventions—that is the direct comparisons. Reviews that intend to compare multiple treatments with an NMA but then find that the expectations or assumptions are violated (eg, underlying assumptions of the method are not met), and hence an NMA is not possible or optimal, are also considered in our definition.

NMA risk of bias assessment

A risk of bias assessment would evaluate limitations in the way in which the NMA analysis was planned, analysed and presented. If these methods are inappropriate, the validity of the findings can be compromised.[38] Our tool aims either/or to assess the biases in the individual results of the NMA, and the authors' conclusions.

Bias in results of an NMA

NMA of effect estimates from primary studies can result in overestimation or underestimation of the effects of specific intervention comparisons.[39 40] For example, Chaimani *et al* conducted a network meta-epidemiological study and found that, in the majority of the 32 networks they analysed, small studies tended to exaggerate the true effect estimate of the intervention, possibly due to small-study effects and publication bias.[41] Our tool will focus on the results of an NMA (eg, network characteristics (including geometry, effect modifiers)).[42] This is the approach taken in tools such as the Cochrane Risk of Bias 2 tool for assessing risk of bias in randomised trials.[43]

Bias in the conclusions of an NMA

Bias may be introduced when interpreting the NMA results to draw conclusions. Conclusions may include 'spin' (eg, biased misrepresentation of the evidence, perhaps to facilitate publication) or (erroneous) misinterpretation of the evidence.[44] Ideally, potential biases identified in the results of the NMA might be addressed appropriately when drawing conclusions. Similarly, a well-conducted systematic review draws conclusions that are appropriate to the included evidence and can therefore be free of bias even when

Continued

Box 1   Continued

the primary studies included in the review have high risk of bias.

Data analysis

Questionnaires that were terminated early, where respondents did not go through all questionnaire pages, were included in analyses, but those that were entirely blank were excluded. We measured the time respondents took to fill in a questionnaire regardless of whether it was complete.

Descriptive statistics were calculated for each closed response question including count, frequency, with denominators taken as the number who provided a response to the question. One researcher coded the open-ended questions independently by identifying themes. Respondents' comments on questions 9 and 10 were merged as they were similar in nature.

Delphi process

Design

Our published methodological review to identify items related to bias in NMAs[14] identified 22 items related to bias in NMAs. These items were reworded into concepts by the steering committee because past Delphi groups focused on the wording of the item, and not its main idea. The concepts and questions about the structure of the tool (domains, signalling questions, rating scales) were entered into a Qualtrics survey platform.

Delphi panellists were asked to rate concepts based on a 5-point Likert scale of importance from 1 (not important—should be dropped as a concept to consider) to 5 (very important—must be included) or unable to score.[20 21] If respondents did not provide a rating, the concept was recorded as missing. Respondents were asked to comment on whether they preferred to modify or reword the concepts.[22] Free-text comment boxes allowed experts to provide additional comments. Non-responders or those failing to complete each round were sent three email reminders.[21]

The respondents completed two survey rounds to reach a high level of agreement, defined as at least 70% scored 4 or above on the 5-point Likert scale[21 23] (table 1). After round 1, we generated reports of group versus individual responses. Respondents were also provided with anonymised free-text comments from the last round.[24]

Data analysis

We reported the number of respondents completing each round. An overall response rate was calculated as well as summary statistics for each concept. The qualitative data from the free-text questions was analysed through thematic analysis by one author (CL) and read by one or more of the other coauthors. The steering committee used their executive power to decide whether concepts excluded by the Delphi panel should be retained or not.

Results

Knowledge user survey

Recruitment results

A total of 2821 emails were sent out to advertise the survey, 87 failed to reach the recipients due to incorrect addresses, resulting in 2734 emails that reached the intended individual (online supplemental appendix H). Most respondents completed the survey through our Qualtrics email survey link (n=390, response

Table 1 Decision criteria for inclusion, exclusion and further consideration of potential concepts*

Scenario (rounds 1 and 2)	Handling of information.
Concept scored 4–5 by ≥70% of respondents	Consensus achieved for potential inclusion in the RoB NMA tool. Further consideration in a subsequent Delphi round not needed.
Concept scored 3 by ≥70% of respondents	Include in Delphi round 2.
Concept scored 1 or 2 (not important) by ≥70% of respondents	Consider excluding the concept.
New concepts nominated by respondents	Include in round 2. Follow decision criteria scenarios above.

NMA, network meta-analysis; RoB, risk of bias.

rate=14%) and 27 completed the survey through an anonymous link distributed over social media and e-newsletters (n=27).

After consolidating duplicates (using IP addresses, n=33) and blank responses (n=86), a total of 298 responses were included in the analysis. Of the 298 respondents, 252 (85%) answered all the survey questions and 46 (15%) completed half of the questions. The mean time to complete the survey was 2.27 min (SD 1.33).

Characteristics of respondents
Of the 298 respondents, 136 (45.6%) self-identified as a systematic review expert, 122 (40.9%) as a guideline developer, 98 (32.9%) as a healthcare professional (table 2). Half of the respondents had primary affiliations at a university (50.0%). Most respondents resided in North America (40.6%) and/or Europe (33.9%) (table 2). Three-quarters (75.1%) of respondents indicated that their organisation produced systematic reviews with NMAs, but only 54.2% of knowledge users said they used an NMA in their work (table 2).

Interest and type of tool preferred
Most knowledge users (84%) reported they would use the RoB NMA tool if they received adequate training on how to use it (figure 1). When asked about their level of interest in our tool, 182/298 (61.1%) had high interest, 53/298 (17.8%) had low interest and only one person had no interest. Many respondents said they would use the RoB NMA tool's bias assessment when conducting an overview of reviews, health technology assessment (HTA) or guideline; and to distinguish between NMAs at higher or lower risk of bias.

When we asked knowledge users about the type of tool that might be useful to them or their organisation, half of the respondents (145/298) reported they preferred a tool to assess *both* the bias in individual NMA results and authors' conclusions (figure 2).

Open-ended questions are summarised in online supplemental appendix I tables 1–4. We also report in online supplemental appendix I table 4 respondents' interest in dissemination and engagement activities. The majority of respondents (153/231; 66%) said they would want to read the final study reports, receive updates (147/231; 64%) and receive training in using the new tool (140/231; 61%).

Delphi survey
Recruitment results
The steering committee invited 53 experts to participate in the Delphi surveys, and 19 emails failed for various reasons, resulting in 37 emails that reached the intended individual. Of these, 28 completed round 1 and 22 completed round 2 (flow chart in online supplemental appendix J). The response rate of panellists participating in our study was 28/37 (75.7%) in round 1 and 22/28 (78.6%) in round 2.

Characteristics of round 1 respondents
Of the 28 round 1 respondents, 15 (53.7%) self-identified as statisticians, 10 (35.7%) as academics, 4 (14.3%) as systematic review specialists or scientists, epidemiologists or graduate students/postdoctoral researchers (table 3). More than half of the respondents had a primary affiliations at a university (68%). Most respondents resided in Europe (53.6%) or North America (39.2%) (table 3). Most (96.4%) respondents indicated that their organisation produced systematic reviews with NMAs.

Rating of concepts
Of the 22 concepts, 7 did not reach consensus in round 2 (indicated in red in table 4). Table 4 lists all concepts in the left column that respondents rated from strongly disagree to strongly agree. The second column indicates whether the concept was included based on 70% of agreement (agree and strongly agree combined). The next columns indicate the number of responses over the denominator (number of people who answered) for each rating, percentage responses and the group median. The list of concepts in table 4 is not intended to be used to assess biases in NMAs, but to inform the development of items to be included in our tool.

Structure of the RoB NMA tool
When asked about the structure of the RoB NMA tool, the majority of respondents agreed that a domain-based structure (25/28; 89.3%) with signalling questions (20/28; 71.4%) was preferred. The domain-based structure would be similar to that used in Cochrane Risk of Bias tool and the ROBIS tool. Signalling questions flag aspects of study design related to the potential for bias and aim to help reviewers judge risk of bias. They also agreed (19/28; 67.9%) that the steering committee should provide guidance on how to produce a risk of bias assessment for NMAs, outcomes within a network or authors' conclusions of NMAs.

When asked about their preference for a tool to assess the risk of bias in NMA results and/or the authors' conclusions, the majority of respondents (15/28; 53.6%) preferred a tool to assess bias in both results and conclusions, one-third (10/28; 35.7%) preferred a tool to assess bias in the results only and a minority (3/28; 10.7%) preferred to assess only the NMA authors' conclusions.

Discussion
A majority of knowledge users responded that they had high interest in the RoB NMA tool if they received adequate training on how to use it and said they would use the tool to distinguish between NMAs at higher or lower risk of bias, and to assess an NMA in an overview of reviews, HTA or guideline. Delphi respondents articulated a clear preference for a tool that is domain-based with signalling questions which would be used to assess biases in the results and the authors' conclusions. Seven out of 22 concepts did not reach consensus by the Delphi group, and these concepts

Table 2 Characteristics of knowledge user respondents and familiarity with NMAs

Characteristics	Overall (n=298)
Primary and current roles*	
Systematic reviewer	136 (45.6%)
Academic	122 (40.9%)
Clinician or healthcare professionals	98 (32.9%)
Graduate student/postdoctoral researcher	60 (25.1%)
Epidemiologist	54 (18.1%)
Guideline developer	44 (14.8%)
Independent researcher	42 (14.1%)
HTA producer or specialist	39 (13.1%)
Statistician	38 (12.8%)
Journal editor	31 (10.4%)
Research support	19 (6.4%)
Decision/Policymaker	9 (3.0%)
Information scientist/Medical librarian	6 (2.0%)
Funding agency representative and clinician	3 (1.0%)
Patient partner	3 (1.0%)
Other (methodologist, non-profit organisation worker, knowledge translation specialist, scientific officer, health economist, etc)	11 (3.7%)
Primary affiliation	
University	149 (50.0%)
Hospital and university hospital	61 (20.5%)
Research institute	25 (8.4%)
Government	19 (6.4%)
Non-profit organisation (eg, NGO, charity)	23 (7.7%)
For-profit private organisation (eg, industry)	10 (3.4%)
Other (eg, clinic, HTA organisation, blood service, independent researcher)	11 (3.7%)
Geographic region	
North America/Central America	121 (40.6%)
Europe	101 (33.9%)
Asia	50 (16.8%)
South America	17 (5.7%)
Africa	2 (0.67%)
Pacific Islands	1 (0.34%)
Australia	4 (1.3%)
Other (ie, Middle East, Oceania)	2 (0.67%)
Does your organisation or institution (or work colleagues) produce systematic reviews with NMA?	
Yes	223 (75.1%)
Unsure	32 (10.8%)
No	42 (14.1%)
Missing	1 (0.3%)
Have you used systematic reviews with NMA as a source of evidence in decision making?	
Yes	193 (65.4%)
No	73 (24.7%)
Unsure	29 (9.83%)
Missing	3 (1.0%)
Have you used a systematic review with NMA in your work?	
Yes	160 (54.2%)
No	160 (54.2%)
Unsure	75 (25.4%)
Missing	3 (1.0%)

Continued

Table 2 Continued

Characteristics	Overall (n=298)
If you have used one or more systematic reviews with NMA in your work, did you use:	
Both individual results of NMAs and conclusions	127 (57.7%)
Individual analysis results from the NMA to draw your own conclusions (eg, pooled effect estimate)	72 (32.7%)
NMA authors' conclusions	21 (9.55%)
Missing	78 (26.2%)

*Percentages add to >100% because respondents could provide more than one response.

HTA, heath technology assessment; NGO, non-governmental organisation; NMA, network meta-analysis.

and accompanying comments will be reviewed and considered by the steering committee for eligibility in the tool. The tool is still under development and the list of concepts is not intended to be used to assess biases in NMAs.

Respondents also indicated the need for guidance on how to use the tool to assess biases in the NMA. These results highlight the necessity for clear and easy to understand elaboration and explanation materials plus training, and perhaps the development of more structured guidelines for reaching domain-based risk of bias judgements (eg, algorithms).[25] Many knowledge users erroneously thought the RoB NMA tool's final assessment would be used in an evaluation of the certainty of the evidence (eg, CINeMA (Confidence in Network Meta-Analysis[26]) or Grading of Recommendations, Assessment, Development and Evaluations[27]), even though we clearly stated that our proposed tool is intended for the assessment of the potential biases in an NMA. Only the quantitative results of an NMA (ie, the analysis) are used in a certainty of the evidence evaluation.

We aimed to engage knowledge users and NMA experts early in the tool development for multiple reasons. Engagement would ensure that our tool will be relevant, useable and accepted.[9] The Delphi expert responses provided us with feedback on which concepts would be most relevant, and the knowledge user responses emphasised the desire for future training[10] and desire to help us with dissemination and communication of findings.[10 11]

Implications of this study

Knowledge user[5 7 28–31] and Delphi[12 32–35] surveys have been successfully used to inform the development of other types of tools, systematic reviews and guidelines. Online surveys have
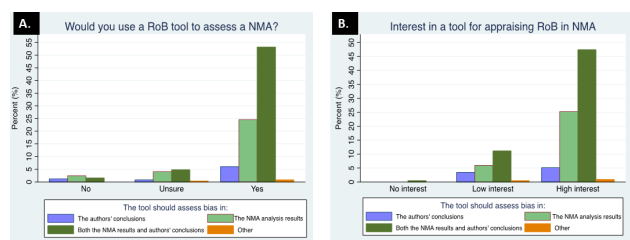

Figure 1 (A) Use of a risk of bias (RoB) network meta-analysis (NMA) tool in knowledge users work and (B) interest in a RoB NMA tool. The left figure (A) depicts the proportion of responses to the question of whether knowledge users would use our proposed RoB NMA tool to assess the NMA analysis results, the authors' conclusions or both results and conclusions. The right figure (B) shows the proportion of responses to the question about interest in a tool for appraising RoB in NMAs.
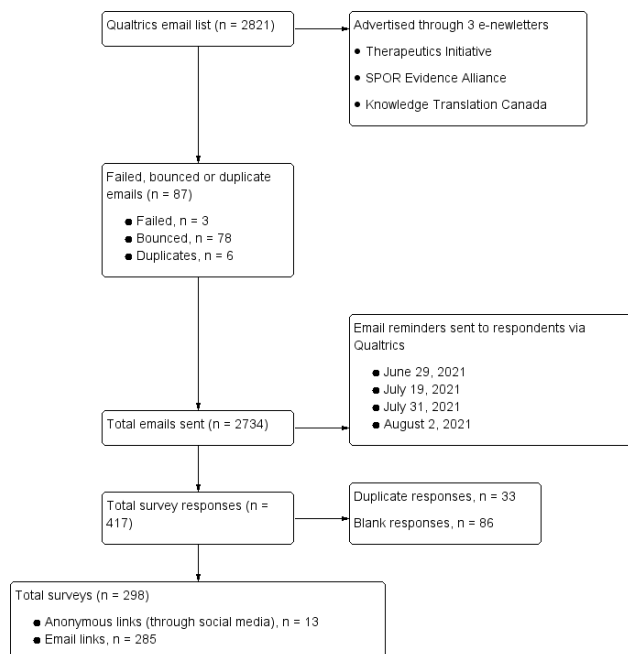
Figure 2 Flow chart of emails sent and responses.

been used to evaluate the reliability and face validity of tools, and the use of the Cochrane risk of bias tool in practice.[25] However, we are not aware of similar published surveys conducted prior to the development of a risk of bias or quality appraisal tool, targeted specifically at knowledge users.

Strengths and limitations

A strength of our research was that we developed a protocol which we followed in the conduct of the research (https://osf.io/da4uy/). We aimed to engage knowledge users and NMA experts early in the tool development for multiple reasons. Engagement would ensure that our tool will be relevant, useable and accepted.[9] The responses provided us with feedback on which concepts are most relevant, training needs[10] and future dissemination of findings. We combined newsletter, email distribution lists and social media to reach a wide range of knowledge users from across the globe. We attempted to maximise the response rate by sending email reminders and repeating messages through social media. Response bias in our sample is a major limitation as knowledge users working in high-income countries were more represented, and respondents (ie, systematic reviewers (45.6%) or academics (40.9%)) may have been more likely to have responded to a survey about a new tool to assess the bias in NMAs.

A limitation is we did not ask knowledge users to define what their role was and whether they considered themselves: (i) decision makers; (ii) purchasers of services/pharma products; (iii) professional service providers; (iv) evidence generators or (v) advocates of health promotion. Another limitation is that our targeted emails and social media advertisement may have missed important knowledge users that use NMAs (eg, members of the Canadian Institutes of Health Research, Drug Safety and Effectiveness Network, Methods and Applications Group for Indirect Comparisons Group).

A strength of the Delphi process was that by performing the surveys online, experts from around the world were able to participate. The response rate of Delphi panellists participating to our study was high in both rounds. All comments were thoroughly read by one of the authors (CL), and part of them were read by

Table 3 Characteristics of Delphi round 1 respondents and expertise in NMAs

Characteristics	Overall (n=28)
Primary and current roles*	
Statistician	15 (53.7%)
Academic	10 (35.7%)
Systematic review specialist or scientist	4 (14.3%)
Epidemiologist	4 (14.3%)
Graduate student/postdoctoral researcher	4 (14.3%)
Clinician or allied health professional	2 (%)
HTA producer or specialist	1 (%)
Health economist	1 (%)
Independent researcher	1 (%)
Journal editor	1 (%)
Geographic region	1 (%)
North America/Central America	11 (39.2%)
Europe	15 (53.6%)
Australasia	3 (%)
Primary affiliation*	
University	19 (67.9%)
Research institute	5 (%)
Hospital	4 (%)
Government	1 (%)
For-profit private organisation (eg, industry)	4 (%)
Non-profit organisation (eg, NGO, charity)	1 (%)
Primary affiliation produces systematic reviews with NMA	27 (%)
Type of experience with NMAs	24 (%)
Subject-matter expert or experienced researcher with knowledge of a variety of evidence synthesis methods and practical experience with systematic reviews with NMA	23 (%)
Author with publications relevant to systematic reviews with NMA	19 (%)
Reader and user of systematic reviews with NMA	17 (%)
Author of a tool for systematic reviews with NMA	12 (%)
Perceived expertise in methods used in the conduct of systematic reviews with NMAs on a scale from 0 to 100, where 0 represents 'beginner' and 100 represents 'expert'	Mean 80.3 (SD 13.9)
Perceived expertise in identification of potential biases in NMAs on a scale from 0 to 100, where 0 represents 'beginner' and 100 represents 'expert'.	Mean 71 (SD 22.7)

*Percentages add to >100% because respondents could provide more than one response.

HTA, heath technology assessment; NGO, non-governmental organisation; NMA, network meta-analysis.

one or more of the other coauthors. In addition to a full feedback report, a summary of the comments in round 1 was also given in the next round to panellists. Delphi processes have many limitations, one of which is that they depend too much on the purposive selection of 'experts' and their availability, thus raising the question of whether all relevant perspectives and scientific disciplines have been taken into consideration. Our Delphi panel consisted of a small double digit number, which may risk collecting certain thought collectives and may be an issue of reliability.[36]

Future research

The results of the survey will inform a new tool to assess biases in NMAs. Our tool is not targeted at authors of NMAs, as it does not

Table 4 Rating of concepts by Delphi respondents in round 1 and 2

Domains	Concepts related to risk of bias in network meta-analysis	Round	Included (70% consensus)	Responses Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree	Unable to Score	Group median (Q1, Q3)
Network characteristics/ geometry	1: Whether all interventions in the network (including comparators) were potentially suitable for all eligible study respondents	1	Y	0/27	0/27	2/27 (7%)	10/27 (37%)	15/27 (56%)	1	5 (4, 5)
	2: Whether any interventions were inappropriately excluded from the network (eg, through eligibility criteria or after seeing the results)	1	Y	0/27	1/27 (4%)	3/27 (11%)	11/27 (41%)	12/27 (44%)	1	4 (4, 5)
	3: Whether importantly different intervention strategies were kept as distinct nodes in the network (ie, whether appropriate groupings were made of interventions—lumping vs splitting)	1	Y	0/28	2/28 (7%)	5/28 (18%)	11/28 (39%)	9/28 (32%)	0	4 (3, 5)
Effect modifiers	4: Whether effect-modifying participant characteristics are sufficiently similar across the whole network	1	Y	0/28	0/28	2/28 (7%)	11/28 (39%)	15/28 (53%)	0	4.5 (4, 5)
	5: Whether outcomes and time points are sufficiently similar across the whole network	1	Y	0/28	0/28	4/28 (14%)	14/28 (50%)	10/28 (36%)	0	4 (4, 5)
	6: Whether study-level risks of bias are sufficiently similar across the whole network	1	N	0/28	9/28 (32%)	5/28 (18%)	8/28 (29%)	5/28 (18%)	0	3 (2, 4)
		2	N	2/22 (9%)	8/22 (36%)	5/22 (23%)	5/22 (23%)	2/22 (9%)	0	3 (2, 4)
	7: Whether other trial characteristics are sufficiently similar across the whole network	1	N	0/24	4/24 (17%)	9/24 (38%)	7/24 (29%)	4/24 (17%)	4	3 (3, 4)
		2	N	0/21	4/21 (19%)	8/21 (38%)	8/21 (38%)	1/21 (5%)	1	3 (3, 4)
Statistical synthesis	8: Whether an appropriate prespecified approach was used in node making	1	Y	1/25 (4%)	1/25 (4%)	5/25 (20%)	13/25 (52%)	5/25 (20%)	3	4 (3, 4)
	9: Whether a process was used to define nodes in the network (eg, undertaken independently by two reviewers, following a preplanned node-making process)	1	N	0/25	2/25 (8%)	11/25 (44%)	9/25 (36%)	3/25 (12%)	3	3 (3, 4)
		2	N	1/20 (5%)	1/20 (5%)	5/20 (25%)	11/20 (55%)	2/20 (10%)	2	4 (3, 4)
	10: Whether effect metric(s) for each outcome (eg, ORs, risk difference) in the network were presented with CIs/ credible intervals	1	N	3/28 (11%)	4/28 (14%)	2/28 (7%)	6/28 (21%)	13/28 (46%)	0	4 (2.25, 5)
		2	N	3/22 (14%)	7/22 (32%)	2/22 (9%)	2/22 (9%)	8/22 (36%)	0	3 (2, 5)
	11: If disconnected networks were connected to perform the analysis, whether methods to do this were appropriate	1	Y	0/25	2/25 (8%)	4/25 (16%)	13/25 (52%)	6/25 (24%)	3	4 (3.5, 4)
	12: Whether methods used to represent multi-arm studies in the dataset and in the analysis are appropriate	1	Y	0/28	4/28 (14%)	3/28 (11%)	10/28 (36%)	11/28 (39%)	0	4 (3.25, 5)
	13: Whether assumptions across the network about homogeneity/heterogeneity of effects within comparisons are appropriate	1	Y	0/27	1/27 (4%)	6/27 (22%)	13/27 (48%)	7/27 (26%)	1	4 (3, 4)
	14: Whether a valid approach was used to determine whether there was conflict between direct and indirect sources of evidence on the same comparisons (often called inconsistency or incoherence)	1	Y	0/28	2/28 (7%)	0/28	8/28 (29%)	18/28 (64%)	0	5 (4, 5)
	15: If inconsistency detected, then whether methods such as re-evaluation of the choice of scale, effect modification and similarity of the contributing randomised controlled trials were investigated	1	N	0/28	4/28 (14%)	5/28 (18%)	14/28 (50%)	5/28 (18%)	0	4 (3, 4)
		2	N	0/22	4/22 (18%)	4/22 (18%)	9/22 (41%)	5/22 (23%)	0	4 (3, 4)
	16: If a Bayesian analysis was conducted, whether the selection of prior distributions was justified	1	Y	0/28	3/28 (11%)	2/28 (7%)	13/28 (46%)	10/28 (36%)	0	4 (4, 5)
	17: Whether the analysis appropriately addressed any differences in effect modifiers across different parts of the network	1	Y	0/28	2/28 (7%)	3/28 (11%)	18/28 (64%)	5/28 (18%)	0	4 (4, 4)

Continued

Table 4 Continued

Domains	Concepts related to risk of bias in network meta-analysis	Round	Included (70% consensus)	Responses					Unable to Score	Group median (Q1, Q3)
				Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree		
	18: Whether there was evidence of conflicting results between direct and indirect evidence (often called inconsistency and incoherence in results)	1	Y	0/28	3/28 (11%)	0/28	10/28 (36%)	14/28 (50%)	0	4 (4, 5)
	19: If there were conflicting results between direct and indirect evidence was this addressed appropriately (eg, meta-regression, data extraction errors, redefining the network)	1	Y	0/27	2/27 (7.4%)	1/27 (4%)	13/27 (48%)	11/27 (41%)	1	4 (4, 5)
	20: Evidence that the statistical model, as it was used to get the key results, was not suitable for the data (eg, from analysis of residuals or information criteria such as DIC)	1	N	0/28	7/28 (25%)	6/28 (21%)	9/28 (32%)	6/28 (21%)	0	3.5 (2.25, 4)
		2	N	0/14	1/21 (5%)	9/21 (43%)	7/21 (33%)	4/21 (19%)	1	4 (3, 4)
	21: Whether sensitivity analyses demonstrate that findings were robust to the statistical model and estimation methods (including prior distributions if Bayesian methods were used)	1	N	0/28	2/28 (7%)	9/28 (32%)	8/28 (29%)	9/28 (32%)	0	4 (3, 4.75)
		2	N	0/22	1/22 (5%)	8/22 (36%)	7/22 (32%)	6/22 (27%)	0	4 (3, 4.75)
	22: Whether limitations at study and outcome level (eg, risk of bias), and at review level (eg, incomplete retrieval of identified research, reporting bias) were discussed	1	Y	0/27	1/27 (4%)	4/27 (15%)	8/27 (30%)	13/27 (48%)	1	4 (4, 5)

*Blank responses and 'unable to score' responses were not counted in the denominator.
DIC, deviance information criterion; N, no; Y, yes.

outline methods that should be followed to conduct an NMA. It is targeted at knowledge users such as healthcare providers, policy-makers, physiotherapists, who want to determine if the results of an NMA can be trusted to be at low risk of bias.

The steering committee will use the results of the knowledge user survey to determine preferences around the tools structure, and the Delphi process to choose and refine the concepts in the tool. Concepts will be reworded into signalling questions and categorised into domains. The new tool will then be pilot tested with different knowledge user groups: patients, healthcare providers and researchers. Further research will involve reliability and validity testing.

Conclusions

The surveys provided feedback from knowledge users and experts on their preferences for the structure, focus and concepts included in a proposed RoB NMA tool, which is under development. Both knowledge users and Delphi panellists preferred a tool to assess both the bias in individual NMA results and authors' conclusions. The majority of knowledge users had high interest in the tool and reported they would use it if they received adequate training.

Author affiliations
[1]Knowledge Translation Program, Li Ka Shing Knowledge Institute, St Michael's Hospital, Toronto, Ontario, Canada
[2]Cochrane Hypertension Review Group, The University of British Columbia, Vancouver, British Columbia, Canada
[3]Epidemiology Division, Dalla Lana School of Public Health and Institute for Health, Management, and Evaluation, University of Toronto, Toronto, Ontario, Canada
[4]Ottawa Hospital Research Institute, Ottawa, Ontario, Canada
[5]School of Epidemiology and Public Health, University of Ottawa, Ottawa, Ontario, Canada
[6]MRC Clinical Trials Unit, University College London, London, UK
[7]Population Health Sciences, NIHR Applied Research Collaboration West (ARC West), University Hospitals Bristol and Weston NHS Foundation Trust, Bristol, UK
[8]McGill University, Montreal, Québec, Canada
[9]University of Toronto, Toronto, Ontario, Canada
[10]Department of Medicine, University of Tornto, Toronto, ON, Canada
[11]Department of Health Product Policy and Standards, World Health Organization, Geneva, Switzerland
[12]Centre for Guidelines, National Institute for Health and Care Excellence (NICE), London, UK
[13]Medical Consultancy Department, Federal Joint Committee - Gemeinsamer Bundesausschuss (G-BA), Berlin, Germany
[14]Simon Fraser University, Burnaby, British Columbia, Canada
[15]Health Service Executive, Research and Evidence, Dublin, UK
[16]Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, BC, Canada

Papakonstantinou, Isabelle Bourtron, Sharon Strauss and Jenn Watt.

ORCID iDs
Carole Lunny http://orcid.org/0000-0002-7825-6765
Brian Hutton http://orcid.org/0000-0001-5662-8647
Jennifer Watt http://orcid.org/0000-0002-5296-6013
Robert C Lorenz http://orcid.org/0000-0003-3286-6792

References
1 Moher D, Schulz KF, Simera I, *et al*. Guidance for developers of health research reporting guidelines. *PLoS Med* 2010;**7**:e1000217.
2 Whiting P, Wolff R, Mallett S, *et al*. A proposed framework for developing quality assessment tools. *Syst Rev* 2017;**6**:204.
3 Bullock A, Morris ZS, Atwell C. Collaboration between health services managers and researchers: making a difference? *J Health Serv Res Policy* 2012;**17 Suppl 2**:2–10.
4 Hanney SR, Gonzalez-Block MA, Buxton MJ, *et al*. The utilisation of health research in policy-making: concepts, examples and methods of assessment. *Health Res Policy Syst* 2003;**1**:1–28.
5 Innvaer S, Vist G, Trommald M, *et al*. Health policy-makers' perceptions of their use of evidence: a systematic review. *J Health Serv Res Policy* 2002;**7**:239–44.
6 Lavis J, Davies H, Oxman A, *et al*. Towards systematic reviews that inform health care management and policy-making. *J Health Serv Res Policy* 2005;**10 Suppl 1**:35–48.
7 Vokó Z*et al*. Similarities and differences between stakeholders' opinions on using HTa information across five European countries. *Improving the Use of an Economic Decision Support Tool* 2018;**14**:149.
8 Vokó Z, Cheung KL, Józwiak-Hagymásy J, *et al*. Similarities and differences between stakeholders' opinions on using Health Technology Assessment (HTA) information across five European countries: results from the EQUIPT survey. *Health Res Policy Syst* 2016;**14**:1–17.
9 Richards C, Carter C, Sherlock K. *Practical approaches to participation*. Macaulay Institute, 2004.
10 Cottrell E, Whitlock E, Kato E, *et al*. *Defining the benefits of stakeholder engagement in systematic reviews*. Rockville, MD: Agency for Healthcare Research and Quality (US), 2014.
11 Deverka PA, Lavallee DC, Desai PJ, *et al*. Stakeholder participation in comparative effectiveness research: defining a framework for effective engagement. *J Comp Eff Res* 2012;**1**:181–94.
12 Whiting P, Savović J, Higgins JPT, *et al*. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol* 2016;**69**:225–34.
13 Shea BJ, Reeves BC, Wells G, *et al*. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017;**358**:j4008.
14 Lunny C, Tricco AC, Veroniki A-A, *et al*. Methodological review to develop a list of bias items used to assess reviews incorporating network meta-analysis: protocol and rationale. *BMJ Open* 2021;**11**:e045987.
15 CIHR. *Knowledge user engagement*. Canadian Institute for Health Research, 2016.
16 Eysenbach G. Improving the quality of web surveys: the checklist for reporting results of Internet E-Surveys (cherries). *J Med Internet Res* 2004;**6**:e34.
17 Gupta K. *A practical guide to needs assessment*. John Wiley & Sons, 2011.
18 Dillman DA. *Mail and Internet surveys*. **18**. 2nd edition. Hoboken, New Jersey: John Wiley & Sons Inc, 2007.
19 Ban JK, Tadrous M, Lu AX, *et al*. History and publication trends in the diffusion and early uptake of indirect comparison meta-analytic methods to study drugs: animated coauthorship networks over time. *BMJ Open* 2018;**8**:e019110.
20 Tricco AC, Zarin W, Antony J, *et al*. An international survey and modified Delphi approach revealed numerous rapid review methods. *J Clin Epidemiol* 2016;**70**:61–7.
21 Pandor A, Kaltenthaler E, Martyn-St James M, *et al*. Delphi consensus reached to produce a decision tool for selecting approaches for rapid reviews (Starr). *J Clin Epidemiol* 2019;**114**:22–9.

22 Page MJ, McKenzie JE, Bossuyt PM, *et al.* Mapping of reporting guidance for systematic reviews and meta-analyses generated a comprehensive item bank for future reporting guidelines. *J Clin Epidemiol* 2020;**118**:60–8.

23 Stevens A*et al. Developing PRISMA-RR, a reporting guideline for rapid reviews of primary studies (protocol.* Oxford, UK: EQUATOR Network, 2018. https://www.equator-network.org/wp-content/uploads/2018/02/PRISMA-RR-protocol.pdf

24 Boulkedid R, Abdoul H, Loustau M, *et al.* Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. *PLoS One* 2011;**6**:e20476.

25 Savović J, Weeks L, Sterne JAC, *et al.* Evaluation of the Cochrane collaboration's tool for assessing the risk of bias in randomized trials: focus groups, online survey, proposed recommendations and their implementation. *Syst Rev* 2014;**3**:37.

26 Nikolakopoulou A, Higgins JPT, Papakonstantinou T, *et al.* Cinema: an approach for assessing confidence in the results of a network meta-analysis. *PLoS Med* 2020;**17**:e1003082.

27 Puhan MA, Schünemann HJ, Murad MH, *et al.* A grade Working group approach for rating the quality of treatment effect estimates from network meta-analysis. *BMJ* 2014;**349**:g5630.

28 Haddaway NR, Kohl C, Rebelo da Silva N, *et al.* A framework for stakeholder engagement during systematic reviews and maps in environmental management. *Environ Evid* 2017;**6**:11.

29 Helbig N, Dawes S, Dzhusupova Z, *et al.* Stakeholder engagement in policy development: observations and lessons from international experience. In: *Policy practice and digital science.* Springer, 2015: 177–204.

30 Sultan S, Siedler MR, Morgan RL, *et al.* An international needs assessment survey of guideline developers demonstrates variability in resources and challenges to collaboration between organizations. *J Gen Intern Med* 2021:1–9.

31 Sibbald SL, Kang H, Graham ID. Collaborative health research partnerships: a survey of researcher and knowledge-user attitudes and perceptions. *Health Res Policy Syst* 2019;**17**:92.

32 Booth A, Clarke M, Ghersi D, *et al.* Establishing a minimum dataset for prospective registration of systematic reviews: an international consultation. *PLoS One* 2011;**6**:e27319.

33 Harrison R, Jones B, Gardner P, *et al.* Quality assessment with diverse studies (QuADS): an appraisal tool for methodological and reporting quality in systematic reviews of mixed- or multi-method studies. *BMC Health Serv Res* 2021;**21**:144.

34 Mokkink LB, Boers M, van der Vleuten CPM, *et al.* COSMIN risk of bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: a Delphi study. *BMC Med Res Methodol* 2020;**20**:1–13.

35 Whiting P, Rutjes AWS, Reitsma JB, *et al.* The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;**3**:1–13.

36 Niederberger M, Spranger J. *Delphi technique in health sciences: a* MAP. *Front Public Health* 2020;**8**:457.

37 Higgins JP, Thomas J, Chandler J, *et al. Cochrane Handbook for systematic reviews of interventions.* John Wiley & Sons, 2019.

38 Savović J, Turner RM, Mawdsley D, *et al.* Association Between Risk-of-Bias Assessments and Results of Randomized Trials in Cochrane Reviews: *The ROBES Meta-Epidemiologic Study. Am J Epidemiol* 2018;**187**:1113–22.

39 Mavridis D, Welton NJ, Sutton A. A selection model for accounting for publication bias in a full network meta-analysis. *Stat Med* 2014;**33**:5399–412.

40 Mavridis D, Efthimiou O, Leucht S, *et al.* Publication bias and small-study effects magnified effectiveness of antipsychotics but their relative ranking remained invariant. *J Clin Epidemiol* 2016;**69**:161–9.

41 Chaimani A, Vasiliadis HS, Pandis N, *et al.* Effects of study precision and risk of bias in networks of interventions: a network meta-epidemiological study. *Int J Epidemiol* 2013;**42**:1120–31.

42 Cote MP, Lubowitz JH, Brand JC, *et al.* Understanding network meta-analysis (NMa) conclusions requires scrutiny of methods and results: introduction to NMa and the geometry of evidence. *Arthroscopy* 2021;**37**:2013–6.

43 Sterne JAC, Savović J, Page MJ, *et al.* Rob 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;**366**:l4898.

44 Boutron I, Ravaud P. Misrepresentation and distortion of research in biomedical literature. *Proc Natl Acad Sci U S A* 2018;**115**:2613–9.

Appendix A: Steering committee of experts in NMA and risk of bias tool development
Appendix B: **Checklist for Reporting Results of Internet E-Surveys**
Appendix C: Knowledge user survey questions
Appendix D: List of organizations and institutions producing NMAs
Appendix E: Email invitation describing the purpose of the knowledge user survey
Appendix F: Knowledge translation plan for dissemination of the knowledge user survey
Appendix G: Knowledge user survey Twitter Campaign
Appendix H: Flowchart of social media ads sent
Appendix I: Summary of results of the open-ended questions from the knowledge user survey
Appendix J: Flowchart from recruitment of the Delphi rounds

Appendix A: Steering committee of experts in NMA and risk of bias tool development

- Lunny C, Cochrane Hypertension Review Group and the Therapeutics Initiative, University of British Columbia, Canada
- Veroniki A, School of Education, University of Ioannina, Ioannina, Greece
- Dias, S, Centre for Reviews and Dissemination, University of York, York, UK
- Hutton, B, Ottawa Hospital Research Institute, Ottawa, Canada. Ottawa University, School of Epidemiology and Public Health, Ottawa, Canada
- Wright J, Cochrane Hypertension Review Group and the Therapeutics Initiative, University of British Columbia, Canada
- White IR, MRC Clinical Trials Unit at UCL, London, UK
- Whiting P, Population Health Sciences, Bristol Medical School, University of Bristol
- Tricco AC, Knowledge Translation Program, Li Ka Shing Knowledge Institute, St. Michael's Hospital, Unity Health Toronto, 209 Victoria Street, East Building, Toronto, ON, M5B 1T8, Canada

Appendix B

Checklist for Reporting Results of Internet E-Surveys (CHERRIES)

Checklist Item	Explanation	Page Number
Describe survey design	Describe target population, sample frame. Is the sample a convenience sample? (In "open" surveys this is most likely.)	Page 8
IRB approval	Mention whether the study has been approved by an IRB.	Page 7
Informed consent	Describe the informed consent process. Where were the participants told the length of time of the survey, which data were stored and where and for how long, who the investigator was, and the purpose of the study?	Page 7
Data protection	If any personal information was collected or stored, describe what mechanisms were used to protect unauthorized access.	Page 7
Development and testing	State how the survey was developed, including whether the usability and technical functionality of the electronic questionnaire had been tested before fielding the questionnaire.	Page 7
Open survey versus closed survey	An "open survey" is a survey open for each visitor of a site, while a closed survey is only open to a sample which the investigator knows (password-protected survey).	Page 8
Contact mode	Indicate whether or not the initial contact with the potential participants was made on the Internet. (Investigators may also send out questionnaires by mail and allow for Web-based data entry.)	Page 8
Advertising the survey	How/where was the survey announced or advertised? Some examples are offline media (newspapers), or online (mailing lists – If yes, which ones?) or banner ads (Where were these banner ads posted and what did they look like?). It is important to know the wording of the announcement as it will heavily influence who chooses to participate. Ideally the survey announcement should be published as an appendix.	Page 7-8
Web/E-mail	State the type of e-survey (eg, one posted on a Web site, or one sent out through e-mail). If it is an e-mail survey, were the responses entered manually into a database, or was there an automatic method for capturing responses?	Page 7-8
Context	Describe the Web site (for mailing list/newsgroup) in which the survey was posted. What is the Web site about, who is visiting it, what are visitors normally looking for? Discuss to what degree the content of the Web site could pre-select the sample or influence the results. For example, a survey about vaccination on a anti-immunization Web site will have different results from a Web survey conducted on a government Web site	NA as we used Qualtrics
Mandatory/voluntary	Was it a mandatory survey to be filled in by every visitor who wanted to enter the Web site, or was it a voluntary survey?	Page 7

Incentives	Were any incentives offered (eg, monetary, prizes, or non-monetary incentives such as an offer to provide the survey results)?	Page 7
Time/Date	In what timeframe were the data collected?	Page 8
Randomization of items or questionnaires	To prevent biases items can be randomized or alternated.	Page 7
Adaptive questioning	Use adaptive questioning (certain items, or only conditionally displayed based on responses to other items) to reduce number and complexity of the questions.	NA
Number of Items	What was the number of questionnaire items per page? The number of items is an important factor for the completion rate.	Page 7
Number of screens (pages)	Over how many pages was the questionnaire distributed? The number of items is an important factor for the completion rate.	Page 7
Completeness check	It is technically possible to do consistency or completeness checks before the questionnaire is submitted. Was this done, and if "yes", how (usually JAVAScript)? An alternative is to check for completeness after the questionnaire has been submitted (and highlight mandatory items). If this has been done, it should be reported. All items should provide a non-response option such as "not applicable" or "rather not say", and selection of one response option should be enforced.	NA Page 7
Review step	State whether respondents were able to review and change their answers (eg, through a Back button or a Review step which displays a summary of the responses and asks the respondents if they are correct).	Page 7
Unique site visitor	If you provide view rates or participation rates, you need to define how you determined a unique visitor. There are different techniques available, based on IP addresses or cookies or both.	NA
View rate (Ratio of unique survey visitors/unique site visitors)	Requires counting unique visitors to the first page of the survey, divided by the number of unique site visitors (not page views!). It is not unusual to have view rates of less than 0.1 % if the survey is voluntary.	NA
Participation rate (Ratio of unique visitors who agreed to participate/unique first survey page visitors)	Count the unique number of people who filled in the first survey page (or agreed to participate, for example by checking a checkbox), divided by visitors who visit the first page of the survey (or the informed consents page, if present). This can also be called "recruitment" rate.	NA
Completion rate (Ratio of users who	The number of people submitting the last questionnaire page, divided by the number of people who agreed to participate (or submitted the first survey page). This is only relevant if there is a separate "informed	NA

finished the survey/users who agreed to participate)	consent" page or if the survey goes over several pages. This is a measure for attrition. Note that "completion" can involve leaving questionnaire items blank. This is not a measure for how completely questionnaires were filled in. (If you need a measure for this, use the word "completeness rate".)	
Cookies used	Indicate whether cookies were used to assign a unique user identifier to each client computer. If so, mention the page on which the cookie was set and read, and how long the cookie was valid. Were duplicate entries avoided by preventing users access to the survey twice; or were duplicate database entries having the same user ID eliminated before analysis? In the latter case, which entries were kept for analysis (eg, the first entry or the most recent)?	NA
IP check	Indicate whether the IP address of the client computer was used to identify potential duplicate entries from the same user. If so, mention the period of time for which no two entries from the same IP address were allowed (eg, 24 hours). Were duplicate entries avoided by preventing users with the same IP address access to the survey twice; or were duplicate database entries having the same IP address within a given period of time eliminated before analysis? If the latter, which entries were kept for analysis (eg, the first entry or the most recent)?	Page 7
Log file analysis	Indicate whether other techniques to analyze the log file for identification of multiple entries were used. If so, please describe.	NA – IP addresses were used
Registration	In "closed" (non-open) surveys, users need to login first and it is easier to prevent duplicate entries from the same user. Describe how this was done. For example, was the survey never displayed a second time once the user had filled it in, or was the username stored together with the survey results and later eliminated? If the latter, which entries were kept for analysis (eg, the first entry or the most recent)?	NA
Handling of incomplete questionnaires	Were only completed questionnaires analyzed? Were questionnaires which terminated early (where, for example, users did not go through all questionnaire pages) also analyzed?	Page 8
Questionnaires submitted with an atypical timestamp	Some investigators may measure the time people needed to fill in a questionnaire and exclude questionnaires that were submitted too soon. Specify the timeframe that was used as a cut-off point, and describe how this point was determined.	Page 8
Statistical correction	Indicate whether any methods such as weighting of items or propensity scores have been used to adjust for the non-representative sample; if so, please describe the methods.	NA

This checklist has been modified from Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res. 2004 Sep 29;6(3):e34 [erratum in J Med Internet Res. 2012; 14(1): e8.]. Article available at

Appendix C_Survey Questionnaire

Introduction

Welcome to the Knowledge User Survey for the Risk of Bias in Network Meta-analysis (RoB NMA) tool project. This study is led by a steering group of international experts in tool development, bias and NMAs including: Drs Julian Higgins, Ian White, Sofia Dias, Argie Veroniki, Andrea Tricco, Penny Whiting, Jim Wright, Brian Hutton, and Carole Lunny.

This survey is part of a larger project to develop a risk of bias assessment tool for network meta-analyses (NMAs). The tool aims to assess bias with a focus on internal validity only: "a systematic error or deviation from the truth, in the summary estimates and/or conclusions". This project is funded by a CIHR project grant (2021-2024).

The purpose of this survey is to ask users of NMAs and knowledge users about what type of tool for assessing an NMA would be most useful. We are also, optionally, giving you the opportunity to look at an initial list of risk of bias items for NMAs in case you think there are any items we may have missed.

Our study protocol and objectives can be found here.

If you have any questions or comments, please contact the principal investigator, Dr Carole Lunny at carole.lunny@ubc.ca

--
Carole Lunny, MPH, PhD
Postdoctoral Fellow, Methodology and Research Synthesis
carole.lunny@ubc.ca
@carole_lunny

Instructions

Your answers to this survey will be used to inform the development of the tool. This survey is voluntary and you may

exit the survey at any time.

Survey Instructions

The survey has 15 questions in total. It should take you 10 minutes to complete sections 1 to 4. Section three is optional and asks you to read the list of items related to bias in NMAs and make suggestions of items not covered, and this may take considerable time depending on how much reflection and work you would like to do. You can skip through any question or section and submit your survey answers on the last page. You can send any comments to the primary investigator, Dr Carole Lunny at carole.lunny@ubc.ca

Data Protection Statement

All data collected in this survey will be stored anonymously and securely. We do not retain any personal data except with your permission. Cookies (i.e. personal data stored by your Web browser) are not used in this survey. We may quote your responses but they will not be attributed to you.

Block 2

Section One: Demographic Information

We would like to start by asking you a few questions about yourself and your work.

1. Indicate your current role (check all that apply)

☐ Health Technology Assessment (HTA) producer or specialist

☐ Funding agency representative

☐ Guideline developer

☐ Decision/policy maker

☐ Epidemiologist

☐ Independent researcher

☐ Academic

☐ Information scientist/medical librarian

☐ Systematic reviewer

☐ Clinician or allied health professional

☐ Statistician

☐ Research support

☐ Graduate student/postdoctoral researcher

☐ Journal editor

☐ Other, please specify: [_____]

2. What is your primary affiliation?

○ University

○ Non-profit organization (e.g., NGO, charity)

○ Research institute

○ University hospital

○ Government

○ Hospital

○ For-profit private organization (e.g. industry)

○ Other, please specify: [   ]

3. In which geographic location do you reside?

○ North America/Central America

○ South America

○ Europe

○ Africa

○ Asia

○ Australia

○ Caribbean Islands

○ Pacific Islands

○ Prefer not to say

○ Other, please specify: [   ]

4. What organization or institution do you work for? Please leave this blank if you would like to remain anonymous.

[   ]

5. Does your organization or institution (or work colleagues) produce systematic reviews with network meta-analysis?

○ Yes

○ No

○ Unsure

6. Have you used systematic reviews with network meta-analyses as a source of evidence in decision making?

○ Yes

○ No

○ Unsure

7. Have you, or your organization/institution, used a systematic reviews with network meta-analysis in your work?

○ No

○ Unsure

○ Yes; if yes, please describe how you used the review:

8. If you have used one or more systematic reviews with NMA in your work, did you use:

○ Individual analysis results from the NMA to draw your own conclusions (e.g. pooled effect estimate from one outcome, rank order of a treatment)

○ The NMA authors' conclusions

○ Used both individual results and conclusions

Block 3

Section Two: Design of the RoB NMA tool
If a review is affected by bias, the results and/or conclusions may be misleading, and not doing a good job of telling the 'truth' about the real difference between an intervention and a control or comparator. In this section we would like to know what sort of tool might be useful to you or your organization.

What type of bias?:
Option A) Bias in results of an individual NMA
Network meta-analysis of effect estimates from primary studies can result in over-estimation or under-estimation of the effects of specific interventions against specific comparators. One option for the tool is to focus on the numeric results of the NMA (including results around intervention rankings). This is the approach taken in tools such as the RoB 2 tool for assessing risk of bias randomized trials and can be accessed here.

Option B) Bias in the conclusions of an NMA
An alternative is to consider bias in the interpretation of the NMA within the wider context. Bias may be introduced when interpreting the NMA results to draw conclusions (for example, conclusions may not be supported by the evidence presented, the relevance of the included primary studies may not have been considered by NMA authors, and reviewers may inappropriately emphasize results on the basis of their statistical significance). Alternatively, potential biases identified in the results of the NMA might be addressed appropriately when drawing conclusions. A well-conducted systematic review draws conclusions that are appropriate to the included evidence and can therefore be free of bias even when the primary studies included in the review have high risk of bias. This is the approach taken in tools such as the ROBIS tool for assessing risk of bias in systematic reviews and can be accessed here.

9. Which option do you prefer?

○ Option A

○ Option B

○ Both option A and B

○ Other, please comment:

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ EBM

10. The RoB NMA tool will be used to assess the methodological features known to increase the risk of bias in the results and/ or the NMA's conclusions. Would you prefer a tool to assess the bias in the individual analysis results of an NMA, or the NMA authors conclusions?

○ Assess risk of bias in the individual analysis results

○ Assess risk bias in the NMA authors conclusions

○ Both individual results and authors conclusions

○ Other (please specify)

11. Would you use a risk of bias tool to assess an NMA (if you had received adequate training on how to use it)?

○ Yes

○ No

○ Unsure

12. If you were to use our new RoB NMA tool, how would you use the results of your risk of bias assessment? (Open ended question)

Block 4

Section Three: Additional NMA bias items (Optional)
This section is optional and asks you to read the list of items related to bias in NMAs and make suggestions of items not covered.

The RoB NMA tool is intended to be used as an extension to the ROBIS tool to assess the risk of bias in systematic reviews. ROBIS (Risk Of Bias In Systematic reviews) is designed to assess the risk of bias in reviews with or without pairwise meta-analysis. The ROBIS tool involves the assessment of methodological features in reviews known to increase the risk of bias categorised into four domains (study eligibility criteria; identification and selection of studies;

data collection and study appraisal; and synthesis and findings). Hence, the items in the RoB NMA tool do not focus on general systematic review methods. For example, an item about the process of selecting studies is not needed as it is about general systematic review methods. Some item concepts might be similar to a ROBIS item but may need additional guidance for NMAs. In this case, we will include the concept.

13. A list of concepts (i.e. items) potentially related to bias in NMAs can be accessed here. Do not be concerned about the wording of the item. We are only concerned about the concept the item conveys.

If you would like to read through the list and make any suggestions for concepts related to biases in NMAs that are important to you but are not covered, please add them in the box here:

Block 5

Section Four: Interest and engagement in development, piloting, dissemination and training

14. How much in interest do you have in a tool for appraising the risk of bias in NMAs?

○ High interest

○ Low interest

○ No interest

https://ubc.ca1.qualtrics.com/Q/EditSection/Blocks/Ajax/GetSurveyPrin...

15. Please indicate your interest in being further engaged in this project (select all that applies):

☐ Being on an email list to receive project updates

☐ Being involved in piloting a new tool to assess the risk of bias in NMAs

☐ Receiving training in using the new tool

☐ Reading the final study reports

☐ Disseminating the research

☐ No interest in being further involved

☐ Other (please specify)

If you are interested in being further involved, please sign up for email updates here.

Appendix D: List of organizations and institutions producing NMAs

Organisation	website
Cochrane Multiple Treatments Methods Group	https://methods.cochrane.org/methods-groups
Campbell Collaboration	https://campbellcollaboration.org
Joanna Briggs Intitute	https://joannabriggs.org
Guidelines International Network	https://g-i-n.net/home
U.S. Agency for Healthcare Research & Quality's Evidence-based Practice Centre program	
Centre for Reviews and Dissemination	https://www.york.ac.uk/crd/
Canadian Agency for Drugs and Technologies in Health	https://www.cadth.ca
Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre)	https://eppi.ioe.ac.uk/cms/
Centre for Implementation Research at the Ottawa Hospital Research Institute	http://www.ohri.ca/cir/
CINeMA – Confidence in network meta-analysis	https://cinema.ispm.unibe.ch/

Appendix E: Email invitation describing the purpose of the knowledge user survey

Dear everyone,

We are launching a knowledge user survey today to ask users of network meta-analyses (NMAs) and knowledge users about what type of tool for assessing the risk of bias in NMAs (RoB NMA tool) would be the most useful. The survey will take approximately *10 minutes* of your time. Everyone is welcome to participate in the survey -- this includes those who know little to none about NMAs as well as experts in the field. We are just trying to get a feel for the interest in our tool.

This survey is part of a larger project to develop a risk of bias assessment tool for reviews with network meta-analysis (RoB NMA tool). The tool aims to assess bias with a focus on internal validity only: "a systematic error or deviation from the truth, in the summary estimates and/or review conclusions".

Follow this link to the Survey:
Take the Survey
Or copy and paste the URL below into your internet browser:
https://ubc.ca1.qualtrics.com/jfe/form/SV_892nehjUTOmXTh4?Q_CHL=email
The survey will be running from today, June 28th, to August 1st, 2021.

Here are a few things you can do to help promote our survey:
•   Forward this email with the survey link to any colleagues you think might be interested
•   Re-tweet/share our survey on your Twitter feed, Facebook page, LinkedIn, and other social media outlets
•   Share the knowledge user survey in your newsletter or internal email list
Thank you for your interest, and any questions or comments can be directed to the principal investigator, Dr. Carole Lunny, at carole.lunny@ubc.ca

Kind regards,
Dr. Carole Lunny and Team

Carole Lunny, MPH, PhD
Postdoctoral Fellow, Methodology and Research Synthesis
carole.lunny@ubc.ca
Twitter: @carole_lunny

Appendix F: Knowledge translation plan for dissemination of the knowledge user survey

Project title	NMA Risk of Bias Tool and surveys
Project description	The risk of bias in NMAs tool is designed to provide knowledge users (i.e., methodologists, statisticians, peer reviewers, professors, guideline developers, policy-makers, researchers more broadly) a framework for assessing risk of bias in network meta-analyses. A proposed supplementary training materials/resource package is intended to build the knowledge needed to use the tool.
What is your overall timeline for dissemination?	2-3 weeks total for the surveys and final tool: • Build anticipation for upcoming tool launch 1 week before the launch; • launch tool; • 1 week post launch of tool
Describe dissemination goal (consult with PI or manager as required)	To make knowledge users aware of the tool and support use of the tool. To support uptake of the tool, training on Risk of Bias in Network Meta-Analysis is needed.
What are your GENERAL key messages?	"We have developed a tool and supporting resources to help users assess risk of bias in network meta-analyses.
What product would you like to share?	Tool & article (article explains the tool) Training resources on: Risk of Bias in NMAs
What would you like your target audience to do with the tool? (e.g., use the tool, read the article, etc.)	Understand how to use the tool Use the tool in their work/projects Disseminate the tool Pilot the tool
Dissemination strategies (e.g., post on website, send through a newsletter, email knowledge users)	- Individual Email informing about the surveys and final tool - Twitter - Newsletters (e.g., KT Canada. Identify relevant newsletters for target audiences) – ADD Organizations Websites announcements (KTP, …) - *Publish tool as a pre-print* - *Seek editorials in journals that publish NMAs* - *Get the tool included in the Cochrane handbook (i.e., chapter on NMAs)* - *Discuss with agencies conducting NMAs to use/cite tool in their guidance*
What materials need to be developed to help with dissemination (e.g., email, newsletter blurb)?	- Individual and organizational emails (mail merge) - Multiple scheduled tweets; twitter cards; targeted hashtags; - Newsletter blurb with possible image - Website announcements (can be tailored from newsletter blurb)

Appendix G: Twitter Campaign for knowledge user survey

General Hashtags:
   #riskofbias
   #NetworkMetaAnalysis
   #kmethodologist
   #statistician
   #healthresearch
   #epitwitter (phase 2 dissemination)
   #medtwitter (phase 2 dissemination)

Hashtag for Tool: #NMARoBTool
*Add hashtags where they fit in naturally within the text, then add additional relevant hashtags at the end of tweet if character count permits (e.g. at end of tweet, append "#cdnhealth #knowledgetranslation")

Twitter Thread (both tweets meet the character count)
First tweet in thread:
knowledge user survey for the RoB NMA tool
@carole_lunny and team are leading the development of a new Risk of Bias for Network Meta-analysis tool (called the RoB NMA tool)! Take part in our knowledge user survey and have your say in how the #NMA tool should be developed! Funded by #CIHR
Link: https://tinyurl.com/tsbr2zcy
1/2"

Second tweet in thread:
"This study is led by Drs. Julian Higgins, Ian White, Sofia Dias, Argie Veroniki, Andrea Tricco, Penny Whiting, Jim Wright, Brian Hutton, and Carole Lunny. This project is funded by a CIHR project grant (2021-2024).
2/2"

Twitter handles:
@kt_program @JennAnnWatt @DrMroz @BCSUPPORTUnit @gba_de @thoefer73 @VMinogue2 @meggomango @bjampoh @ubc @ubcnews @ubc @UofT @CADTH_ACMTS @NICE_DSU @naci @ATricco @sdias_stats @AVeroniki @jmwright4 @Geointheworld @BH_epistat @tweetastevens @bobnakagawa @lorenzomoja @cochranecollab @LucyHenryOtt @WHO @HTAiOrg @Drug_Evidence @SPORAlliance

@cochrane_US @mjpages @METRICStanford @MetaEvidence @CampbellUKIRE @JBI_EI @HEI_mcmaster @GRADE_McMaster @OttMethodsCentr @metaEvidenceOrg @methodsctr @METRIC_Berlin @KSR_SysRev @SysReviews @EPPIReviewer @rapidreviews_i @campbellreviews @CochraneHTN

@cochranemthds @James_M_Thomas @OttMethodsCentr @CochraneCanada @cochranecollab @CochraneRRMG @CochraneSGMG

Appendix H: Flowchart of social media ads sent

Six tweets were sent out which resulted in 28 retweets and 4 comments. One LinkedIn advertisement was sent out.

Flowchart: Social media advertisements to advertise the survey
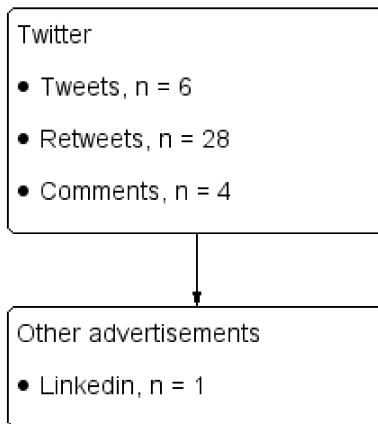
Twitter

- Tweets, n = 6
- Retweets, n = 28
- Comments, n = 4

Other advertisements

- Linkedin, n = 1

Appendix I: Summary of results of the open-ended questions from the knowledge user survey

Appendix Table 1: How NMAs are used in a knowledge users work

How NMAs are used in a knowledge users' work	Frequency	Quotes
Produced NMAs	31	"My centre regularly conducts SRs, including NMA, IPD."
Inform clinical decision making	35	"Have read few NMAs related to my clinical practice and implemented evidence-based on that" "Majorly for finding the effective and tolerable treatments for diseases, when there are no head to head data available"
Inform guideline development, HTAs, or policy decisions	52	"In some of the systematic reviews my organisation does to inform clinical practice guidelines, we have used existing systematic reviews with network meta-analyses to report the quantitative benefits of interventions as well as to determine GRADE ratings for the certainty of the evidence. " "For drug approval and reimbursement and to inform benefit-risk, medical strategy, and the scientific narrative." "NMA have been used for decision-making, to incorporate health technologies in the XXX public health system" "We evaluate medicines candidate to the WHO Model List of Essential Medicines. Some medicine dossiers are largely based on results of NMA."
Inform academic research	35	"Developing background for grant applications and manuscripts; support when writing SR and NMA protocols; conducting SR with NMA" "Discussion and presentation in a journal club for clinicians" "Exploration of evidence-base to inform further studies"
Teaching	5	"Class discussions, for class presentation, in Critical Appraisal exercises and Tutorials. I teach the EBM component of the MMED 1 Curriculum"
Included and used in an 'overviews of reviews'	4	"Consideration in reviews of reviews on tobacco cessation," "If conducting a review of reviews, existing SRs and NMAs would be included by our protocol and incorporated into the synthesis."
Update the registries, databases or websites	2	"XXX is a registry of quality-appraised systematic reviews on the effectiveness and cost-effectiveness of public health interventions/policies. We regularly update our database with new reviews and have been including more and more NMAs."
Economic modelling	5	"NMAs are part of sponsor submissions to the Common Drug Review or Oncology Drug Review process. They are used in economic modelling." "We conducted NMAs to identify likely most effective technologies and to input into cost-utility analyses. We used data from published NMAs to inform economic model parameters."

*Numbers do not add up because participants could provide more than one response.

Comments on whether bias in the individual results of NMA, or authors conclusions were preferred
When asked to comment on whether they preferred a tool to (i) assess bias in the individual results of NMA, or (ii) authors conclusions, 19 out of 249 people commented. Comments in **Appendix Table 2** show that most participants believe both are important.

Appendix Table 2: Knowledge user ' belief in the importance of a tool to assess bias in the individual results of NMA (option A), or authors conclusions (option B) (n = 249)

Themes	Frequency	Quotes
Bias in the individual results of NMA is more important	7	"Option B is superior to option A- as the context of the information is important. However, if seeking the least biased evidence to a clinical question, we also over promote individual studies (within the larger context of the body of evidence) that are from primary care (our home) represent our patients and have meaningful outcomes" "we would tend only to make use of the results and not the interpretation (and combine that with our own assessment of the bias/applicability of the studies) and therefore option B is less directly relevant to the work we do."
Applicability (external validity) should be assessed	2	"Bias should be evaluated with respect to both internal validity and applicability (external validity) to the review question."
Both important	11	"It would depend on the research question I was dealing with. There may be situations when I want to consider the components of an NMA, and other situations when I want to consider NMA as a whole " "Personally I would probably err towards A, i.e. assessing the risk of bias in the statistical results produced by the NMA. However an assessment of bias in the conclusion is also important a) for its own sake and b) as an indicator of the reliability of the NMA as a whole. So I would probably go for for maybe 60% of the questions focusing on A, and 40% on B." "Both option A and B depending on the objective (relating to previous question): Option A if specific results are of interest, Option B if conclusions are relevant or in meta-research. Maybe the tool could consist of both elements and require the researcher to make the aim of the use of the NMA RoB tool transparent (evaluate results, conclusions, or both)." "There are circumstances were option B would have value (for example, when a paper is being peer-reviewed these sorts of considerations would be important)."
Authors conclusions not useful for clinical practice	1	"I think both have merits but option B is not super useful in practice." "I strongly oppose relying on the conclusions of the NMA itself in anyway. There are so many inexperienced individuals conducting NMA and such a horrendous peer review process (assuming that this is due to limited qualified individuals to perform peer review) that I would strongly urge you not to pursue option B which is likely to be of limited value."

*Numbers do not add up because participants could provide more than one response.

Appendix Table 3: Knowledge user' use of a completed NMA risk of bias assessment (n = 145)

Themes	Frequency	Quotes
Inform a policy brief, HTA, clinical practice guideline, or other policy related documents.	22	"This can help methodologist supporting decision-makers to document the risk of bias of selected literature informing a policy brief, or other policy related documents."

To distinguish between NMAs at high or low risk of bias	24	"When using multiple systematic reviews including both pairwise or NMA, the ones using NMA will also be evaluated for risk of bias, and perhaps, determine the most appropriate NMA to inform a particular decision." "Using the RoB NMA tool would reinforce my confidence in NMA results, being these at high or low risk of bias. This would increase the independence of my conclusions from authors' conclusions - basically I will have more chance to form my opinion as a second opinion."
Help producers of NMAs identify issues that may introduce bias (e.g. written in the limitations section of an NMA or a protocol)	16	"It can serve researchers conducting systematic reviews with NMA to be explicit about issues that can introduce bias and avoid them or address them accordingly." "I would discuss them within/after limitations of a NMA." "I would use it to guide the design, implementation and interpretation of my own NMAs"
Help in clinical decision making	1	It can help clinicians who are interested in reading systematic reviews including NMA, by giving them orientation as to what are the factors to consider.
Conduct sensitivity, subgroup or meta-regression analysis	12	"I would use the results of RoB assessments to perform sensitivity analyses (e.g., excluding studies with high RoB), and to determine the certainty of evidence." "We can use this to run some sensitivity analyses or meta-regression to check robustness. Furthermore, in case of high biasness, it could be added as one of the limitations for the review"
Used in the production of 'overviews of reviews'	12	"We often conduct overviews of reviews, therefore a critical appraisal of the included systematic reviews is necessary. Unfortunately, the actual tools for assessing the quality or risk of bias of the NMAs are not satisfying. We would use the RoB NMA tool similarly to NMA as we use the tools ROBIS or AMSTAR2 for the critical appraisal of systematic reviews."
When assessing the certainty in the body of evidence (e.g. using CINeMA, GRADE, or threshold approach)	32	"In assessment of the certainty of the evidence. I would hope that explanations on how to use this together with (or if possible/applicable, in replacement of) individual study ROB results would be available." "We would use it to inform the RoB assessments in the NMAs to judge the certainty in the evidence" "We would use the resalts of the ROB assessment in judging ROB in GRADEpro"
Integrated into the results, interpretation, and conclusions of an NMA, overview of reviews, HTA, or guideline	26	"Draw my own conclusions based on potential for RoB in individual analyses." "I would use it as is recommended in the Cochrane systematic review guidance. Specifically, I would include the result when reporting the results and translate it to 'quality of evidence'."
Peer reviewing or editorial decisions	4	"Peer review"; "Perhaps as co-author or reviewer of a paper that uses NMA"
To inform future research	8	"For future reviews, to know how to consider the results"
In academic research	7	"For methodological research, to compare NMAs" "In conducting our own research, including the development of future research questions, and in writing discussions in our own papers." "Perhaps considering the results of meta-epidemiologic analyses to establish priors that minimize the effect of bias in the ranking"

When teaching research methods	1	"To inform editorial decisions; when teaching research methods and interpretation; when designing SR+NMA protocols"

*Numbers do not add up to the respondent's comments because participants could provide more than one response.

Appendix Table 4: Interest and engagement activities (n= 231)

Activities	Frequency
Being involved in piloting a new tool to assess the risk of bias in NMAs	115
Disseminating the research	70
Reading the final study reports	153
Receiving training in using the new tool	140
Being on an email list to receive project updates	147
No interest in being further involved	19
Other	11

*Numbers do not add up to the respondent's comments because participants could provide more than one response.

Appendix J: Flowchart from recruitment of the Delphi rounds

Round 1

Round 2

Round 1 email invitation sent June 30, 2021
n = 56 emails

- Declined to participate, n = 3
  - Emails failed, n = 9
  - Out of office, n = 7

Round 1 survey email successfully received
n = 37

Completed Round 1
n = 28

First email invitation sent July 22, 2021
n = 26

Four reminders sent
- July 29, 2021
- August 4, 2021
- August 5, 2021
- August 7, 2021

Three reminders sent
- July 7, 2021
- July 14, 2021
- July 16, 2021

Completed Round 2
n = 22