



# When to replicate systematic reviews of interventions: consensus checklist

Peter Tugwell,<sup>1,2,3,4</sup> Vivian Andrea Welch,<sup>2,4</sup> Sathya Karunanathan,<sup>3</sup> Lara J Maxwell,<sup>3</sup> Elie A Akl,<sup>5</sup> Marc T Avey,<sup>6</sup> Zulfiqar A Bhutta,<sup>7</sup> Melissa C Brouwers,<sup>4</sup> Jocalyn P Clark,<sup>8</sup> Sophie Cook,<sup>9</sup> Luis Gabriel Cuervo,<sup>10</sup> Janet Agnes Curran,<sup>11</sup> Elizabeth Tanjong Ghogomu,<sup>2</sup> Ian G Graham,<sup>3,4</sup> Jeremy M Grimshaw,<sup>1,3</sup> Brian Hutton,<sup>3</sup> John P A Ioannidis,<sup>12</sup> Zoe Jordan,<sup>13</sup> Janet Elizabeth Jull,<sup>14</sup> Elizabeth Kristjansson,<sup>15</sup> Etienne V Langlois,<sup>16</sup> Julian Little,<sup>4</sup> Anne Lyddiatt,<sup>17</sup> Janet E Martin,<sup>18</sup> Ana Marušić,<sup>19</sup> Lawrence Mbuagbaw,<sup>20</sup> David Moher,<sup>3</sup> Rachael L Morton,<sup>21</sup> Mona Nasser,<sup>22</sup> Matthew J Page,<sup>23</sup> Jordi Pardo Pardo,<sup>24</sup> Jennifer Petkovic,<sup>2</sup> Mark Petticrew,<sup>25</sup> Terri Pigott,<sup>26</sup> Kevin Pottie,<sup>27</sup> Gabriel Rada,<sup>28</sup> Tamara Rader,<sup>29</sup> Alison Y Riddle,<sup>2</sup> Hannah Rothstein,<sup>30</sup> Holger J Schüneman,<sup>31</sup> Larissa Shamseer,<sup>3,4</sup> Beverley J Shea,<sup>4</sup> Rosiane Simeon,<sup>32</sup> Konstantinos C Siontis,<sup>33</sup> Maureen Smith,<sup>34</sup> Karla Soares-Weiser,<sup>35</sup> Kednapa Thavorn,<sup>3,4</sup> David Tovey Thavorn,<sup>35</sup> Brigitte Vachon,<sup>36</sup> Jeffery Valentine,<sup>37</sup> Rebecca Villemaire,<sup>38</sup> Peter Walker,<sup>2</sup> Laura Weeks,<sup>29</sup> George Wells,<sup>39</sup> David B Wilson,<sup>40</sup> Howard White<sup>41</sup>

For numbered affiliations see end of the article.

Correspondence to: P Tugwell  
ptugwell@uottawa.ca

(ORCID 0000-0001-5062-0556)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2020;370:m2864  
<http://dx.doi.org/10.1136/bmj.m2864>

Accepted: 9 June 2020

Replication is an essential part of the scientific method, yet replication of systematic reviews is too often overlooked, and done unnecessarily or poorly. Excessive replication (doing the same study repeatedly) is unethical and a cause of research wastage. This article provides consensus based guidance on when to replicate and not replicate systematic reviews.

The inability to replicate research findings in various scientific disciplines has been a concern.<sup>1,2</sup> Systematic

reviews are increasingly used as the basis for informing health policy and clinical practice decisions.<sup>3</sup> However, reliance on systematic reviews for decision making assumes that review findings would not change if the review was replicated (eg, methods for the review are repeated by an independent team of reviewers). Given insufficient capacity to replicate all systematic reviews, researchers should determine which reviews warrant replication.

The need for replication of systematic reviews arises from concerns or lack of clarity about the technical or statistical methods or the judgments made, such as the subjective decisions related to defining criteria for inclusion of the population, intervention or exposure, and outcomes of interest; and data collection, synthesis, and interpretation.<sup>4</sup>

The compelling scientific case for replicating systematic reviews is complicated by research waste concerns (eg, multiple reviews of the same question outnumbering the original studies).<sup>5,6</sup> The term “duplication” has been used to refer to needless, frequently unwitting or unacknowledged repetition without a clearly defined purpose for the repetition.<sup>7,8</sup> With finite research resources available and the risk of potentially misleading results being introduced due to unrecognised bias, the issues surrounding replication are intensified.<sup>9-11</sup>

In this article, we discuss criteria for applying replication appropriately. Recognising that the terminology and conceptual framework for replication is not standardised across the sciences, we define replication of systematic reviews of interventions as either:

- Direct replication by purposeful repetition to verify the findings of the original research question; or
- Conceptual replication by purposeful broadening or narrowing of the research question in existing systematic reviews (eg, across broader or more focused populations, intervention types, settings, outcomes, or study designs).

## SUMMARY POINTS

For systematic reviews of interventions, replication is defined as the reproduction of findings of previous systematic reviews looking at the same effectiveness question either by: purposefully repeating the same methods to verify one or more empirical findings; or purposefully extending or narrowing the systematic review to a broader or more focused question (eg, across broader or more focused populations, intervention types, settings, outcomes, or study designs)

Although systematic reviews are often used as the basis for informing policy and practice decisions, little evidence has been published so far on whether replication of systematic reviews is worthwhile

Replication of existing systematic reviews cannot be done for all topics; any unnecessary or poorly conducted replication contributes to research waste

The decision to replicate a systematic review should be based on the priority of the research question; the likelihood that a replication will resolve uncertainties, controversies, or the need for additional evidence; the magnitude of the benefit or harm of implementing findings of a replication; and the opportunity cost of the replication

Systematic review authors, commissioners, funders, and other users (including clinicians, patients, and representatives from policy making organisations) can use the guidance and checklist proposed here to assess the need for a replication

Both types of replication could include differences in systematic review methods (appendix box 1), which include, for example, whether to use traditional comprehensive reviews versus rapid reviews or living systematic reviews. They can involve either independently conducting an entire review or simply repeating a selected part, such as a reanalysis or subgroup analysis.

Systematic review updates and replications are overlapping, but not mutually inclusive.<sup>12</sup> The decision to conduct an update of a systematic review is largely driven by the availability of new data to answer the question of interest, and could occasionally be driven by the availability of new systematic review methods. An update including more data will usually increase the statistical precision of effect estimates. The update could, however, fail to resolve important differences related to previous reviews' technical procedures or subjective decisions, which is best protected against by independent replication.

Although criteria for when to update systematic reviews have been proposed,<sup>13 14</sup> systematic review organisations and research funders or commissioners lack tools to judge the need for replicating previous systematic reviews, and thus have no way of knowing whether replication should be prioritised or when such replications would be a poor use of resources or even waste.

We have developed a checklist and guidance to help decide when to replicate or not replicate systematic reviews of studies looking at the effects of interventions to improve health and wellbeing. The checklist can be used by systematic review authors and organisations, funders, commissioners, and groups developing guidelines, consumer decision aids, and policy briefs.

## Methods

Our methods are adapted from guidance for developing research reporting guidelines.<sup>15</sup> We formed an executive group (SK, LM, JP, JPP, PT, VAW) to gather opinions of stakeholders, review the literature, hold a face-to-face consensus meeting of key stakeholders, and draft the checklist and article (fig 1).

We used an integrated approach for knowledge translation (as previously described<sup>16</sup>), with an international multidisciplinary team of methodologists and knowledge users at every stage of this research. The methodologists were experts in clinical epidemiology, consensus methods, guideline development, health economics, information management, information science, knowledge translation, preclinical research, qualitative methods, and statistics. The knowledge users included commissioners, funders, and other users of systematic reviews, including clinicians, patients, and representatives from organisations involved with policy making. This team was involved in the planning of this project and contributed to the interviews, online survey, consensus meeting, this guidance, as well as the dissemination of the output.

Following empirical and opinion gathering steps (appendices 2-5), the executive group collected data

and developed a 12 item list of criteria for replication (appendix 6), for deliberation at a consensus meeting (6-8 February, 2019) in Wakefield, Canada. During the meeting, each candidate criterion was presented, along with relevant data from the key informant interviews, literature review, and online survey. One designated meeting participant provided a brief commentary on the criterion and two others facilitated an interactive discussion on its value in determining when to replicate systematic reviews. At the end of a 30 minute discussion for each criterion, participants were asked whether they supported the usefulness of an item when deciding when to replicate systematic reviews. Comments were transcribed and considered in subsequent iterations of the checklist.

Following the meeting, the executive group formed the writing group. They summarised the meeting discussions, revised the checklist, and drafted this guidance document. They sought feedback on the checklist and guidance document from all team members through phone meetings and two rounds of manuscript revision. The final checklist was agreed on by all participants through email.

## Patient and public involvement

Seventeen members of the Cochrane consumer group (that is, patients, their families and carers, and members of the public) were recruited as research participants to complete the online survey. Our core research team included two patient partners who are included as authors. They provided input on the grant application, participated in the key informant interviews, online survey and consensus meeting, reflected the voices of consumers who responded to the survey, and revised and approved the manuscript.

## Results

The consensus meeting was attended in person by 36 of the 54 research team members who were invited. They represented a range of disciplines and stakeholder groups. Extensive discussion and debate was conducted in relation to the overall definitions and approaches to replication, the potential use of the checklist, and each of its proposed items. At the end of the meeting, the executive group summarised the main points and provided an opportunity for participants to comment before preliminary agreement on the direction of the checklist was reached.

Participants at the consensus meeting agreed that every systematic review of a policy or practice relevant question should be seriously considered for replication. However, given limited resources, before deciding whether to replicate a systematic review, all existing systematic reviews and overviews of systematic reviews on the same question should be identified, and the added value of conducting a replication assessed. Given the absence of a robust evidence base to justify systematic review replication, the group agreed that a checklist would encourage its users to give due consideration to the pragmatic use

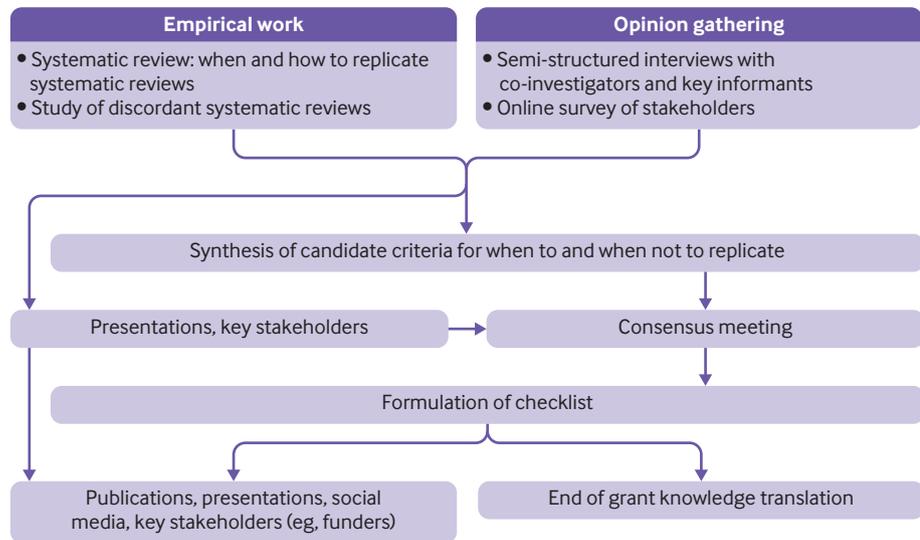


Fig 1 | Overall research plan: checklist development and dissemination

(clinical/policy/fiscal value) of conducting a project of this nature.

Since many of the discussions before the meeting focused on the implicit trade-offs between the benefits and costs of systematic review replication, participants agreed that the value of information (VOI) framework<sup>17 18</sup> could be helpful in developing criteria for when to replicate and when not to replicate prior systematic reviews. VOI is a quantitative approach that calculates the return on investment of conducting additional research (in this case, a replication of a systematic review) to generate clearer evidence to inform a clinical or policy decision, where decision uncertainty currently exists. Most VOI studies use cost effectiveness models to quantify the value of additional research. Based on the consensus meeting, the checklist adopted a conceptual VOI framework<sup>18</sup> and defined the value of systematic review replication as priority (question 1), potential implications (question 2), and potential impacts (question 3). Question 4 reflects the cost of a systematic review replication.

Participants at the consensus meeting decided collectively that the checklist needed to be a high level, simple tool that linked to available detailed tools for each step. For example, participants endorsed a general item on the methodological quality of previous systematic reviews; some quality assessment tools were suggested, such as AMSTAR-2,<sup>19</sup> which provide detailed guidance for users.

Participants also agreed that it would be important to consider the checklist items as examples of reasons for replicating rather than defining an exhaustive list, because some nuances in reasons for replication might be difficult to predict (eg, regarding concerns about conflicts of interest). Finally, participants concurred that replication methods would be driven by uncertainties expressed by users of the reviews (eg, patients and public, practitioners, press, professionals, policy makers). The specific uncertainty

would drive the selection of type of replication—direct replication by purposeful repetition or conceptual replication by purposeful broadening or narrowing, using permutations of previous review methods, data, or assumptions.

### Checklist

After the consensus meeting, the executive group formed the writing group. Members met regularly to revise the checklist on the basis of recommendations emerging from the meeting. The 12 item checklist was concentrated into four items. Some of the 12 items were subsumed under the four broader items. Other items were removed because they appear in the accompanying tools. For example, although the research team generally agreed that the durability of information produced through the replication was worth mentioning, this would generally be part of the item assessing the priority of a systematic review replication, and is included in priority setting tools such as the SPARK tool.<sup>20</sup>

Table 1 presents the final checklist for when a systematic review should be replicated. Checklist items are presented as questions to guide users. The checklist is aimed at systematic review authors and organisations, funders, commissioners, and groups developing guidelines, decision aids, and policy briefs; or anyone using systematic reviews of interventions as the basis of decision making. The extension statement (appendix 7) provides an explanation and example for each item.

### Application of the systematic review replication checklist: example

The checklist was applied to two systematic reviews that were considered under the Campbell International Development Coordinating Group (table 2). The first review was a replication of a 2015 Cochrane review on deworming to improve developmental health and

**Table 1 | Systematic review replication checklist**

Checklist items	Supporting methods and tools	Response options
Question 1. Has the priority for replication been assessed as high? For example, is it likely that a replication will remain relevant to policy and practice for a useful length of time? Is it likely for replication results to lead to implementation by practitioners and policy makers?	Priority setting tool (eg, SPARK tool, <sup>20</sup> James Lind Alliance, <sup>21</sup> CINARI <sup>22</sup> )	Yes / No
Question 2. Is it likely that direct replication by repetition or conceptual replication by broadening or narrowing of the scope will address uncertainties, controversies, or the need for additional evidence related to:		
2.1. The framing of the question in previous reviews?	Stakeholder engagement: James Lind	Yes / No
2.2. The conduct and reporting of previous reviews?	Alliance Priority Setting Partnerships <sup>21</sup> ;	Yes / No
2.3. Author influence or conflicts of interest in previous reviews?	conduct and reporting: AMSTAR-2, <sup>19</sup>	Yes / No
2.4. Discordant findings in previous reviews?	ROBIS <sup>23</sup>	Yes / No
Question 3. Would the implementation of the findings of a replication be likely to have a potentially important sizeable individual benefit or harm or affect a sizeable population?	—	Yes / No
Question 4. Are resources (time, money) best spent on replication rather than on alternative systematic reviews (considering opportunity cost)?	Stakeholder engagement priority setting	Yes / No

wellbeing,<sup>27</sup> and the second was an update of a 2012 Cochrane review on strategies to increase ownership of bednets to prevent malaria.<sup>28</sup> Both reviews involved simple, discrete, low cost interventions implemented by non-governmental organisations to manage a public health issue, with claims of a wide range of benefits. Responses to the checklist provided a strong justification for a replication of the first systematic review on deworming. A replication of the second review would be of much less value, because no substantive discordance was found and therefore would be less controversial.

**Systematic review replication worksheet**

Table 3 presents a replication comparative worksheet for appraising the body of evidence that can include multiple systematic reviews. The worksheet was developed by team members with expertise in guideline development, and it provides an example of

how the checklist can be adapted for use by decision makers. It delineates the decision making process of replication and proposes tools and instruments for each step, although some of these tools remain to be developed. The starting point is a specific question (known as the index PICO (population/intervention/comparison/outcome)) raised by a key stakeholder (eg, guideline panel). The PICOs of existing systematic reviews are then compared to the index PICO to assess their relevance. Other sources of uncertainty related to the conduct or reporting of existing reviews or conflicts of interest are also identified before determining whether a replication could reduce uncertainty related to the index PICO. The final decision to replicate a systematic review or not is based explicit consideration of benefits, harms, cost, feasibility, and acceptability of a given replication against other alternatives. We have plans for validating this worksheet in pilot studies with Cochrane and the Campbell Collaboration.

**Table 2 | Application of systematic review replication checklist on two example reviews**

Checklist items	Review 1: effects of mass deworming	Review 2: strategies to increase ownership and use of bednets
Question 1. Has the priority for replication been assessed as high using appropriate tools for setting research priorities?		
Is it likely for replication results to lead to implementation by practitioners and policy makers?	Yes	No
Is it likely that a replication will remain relevant to policy and practice for a useful length of time?	Relevant to Canadian Red Cross, WHO (update to deworming guidelines planned)	Not ranked as high priority as not being currently debated
Question 2. Is it likely that direct replication by repetition or conceptual replication by broadening or narrowing of the scope will address uncertainties, controversies, or the need for additional evidence related to:	Published critiques of the original review currently exist (long term studies excluded, data cut up, discrepancies with review by Hall et al <sup>24</sup> ), political controversy related to a replication of primary study by Miguel et al (impact of errors) <sup>25 26</sup>	No uncertainties or controversies had been raised
2.1. The framing of the question in previous reviews?	Yes	No
2.2. The conduct and reporting of previous reviews?	Yes	No
2.3. Author influence or conflicts of interest in previous reviews?	No	No
2.4. Discordant findings in previous reviews?	Yes	No
Question 3. Would the implementation of the findings of a replication be likely to have a potentially important sizeable benefit or harm or affect a sizeable population?	Yes, replication would have an important benefit for children in soil transmitted helminth endemic areas. Deworming is part of the implementation programmes for non-governmental organisations, with interest from WHO in updating its guidelines in 2016. Savings from ending deworming programmes could be transferred to education	No, because no change in findings or practice is expected; so no strong case for a sizeable benefit or harm
Question 4. Are resources (time, money) best spent on replication rather than on alternative systematic reviews (opportunity cost)?*	Resources are justified by the potential benefits of resolving the uncertainties around deworming	Although the cost of this update is much less than the deworming replication, the additional knowledge gained (that is, benefits) is also negligible

\*Costs of the exercise should be considered if difference in benefits are marginal.

Table 3 | Comparative worksheet for systematic review replication\*

Step	1. Has the priority for replication been assessed as high?	2.1. Is the framing of the SR PICO not relevant enough to the PICO?	2.2. Is there poor conduct or reporting of the SR?	2.3. Are there concerns about the author influences or conflicts of interests?	2.4. Are there discordant findings across systems reviews addressing the same question?	Overall assessment: Is it likely that replication will address uncertainties, controversies or the need for additional evidence?	Trade-offs of replication assessed: 3. Would the implementation of the findings of a replication be likely to have a potentially important sizeable benefit or harm or affect a sizeable population? 4. Are resources (time, money) best spent on replication rather than alternative high priority SRs?	Decision whether to replicate
Sample tools and resources	Priority setting tools (eg, SPARK tool <sup>(26)</sup> )	GRADE relevance tool (under development), PICO mapping	AMSTAR-2, <sup>19</sup> ROBIS <sup>23</sup>	Conflict of interest funding verification	Jadad algorithm, <sup>29</sup> Moja et al methods, <sup>30</sup> discordance checklist	Based on answers to 2.1-2.4 (if "yes" to any of the questions)	Value of information framework	<b>Judgment of SR team and decision makers based on previous steps</b>
Evaluation of existing SRs and value of SR replication	SR1 (PICO1) SR2 (PICO2) SR3 (PICO3) SR4 (PICO4) SR5 (PICO5) SR6 (PICO6) SR... (PICO...)	Yes / No Yes / No Yes / No Yes / No Yes / No Yes / No Yes / No	Yes / No Yes / No Yes / No Yes / No Yes / No Yes / No Yes / No	Yes / No Yes / No Yes / No Yes / No Yes / No Yes / No Yes / No	Yes / No Yes / No Yes / No Yes / No Yes / No Yes / No Yes / No	Yes / No Yes / No Yes / No Yes / No Yes / No Yes / No Yes / No	Yes / No Yes / No Yes / No Yes / No Yes / No Yes / No Yes / No	Option 1: do not replicate, and use an existing SR. Option 2: partial replication (use existing SR after some fixing) or full replication (direct replication by purposeful repetition; conceptual replication by purposeful broadening or narrowing of the scope)

SR: systematic review; PICO=population/intervention/comparison/outcome.

**Discussion**

Replication is a cornerstone of the scientific method, and critical to best practices; transparency; and accuracy of health, healthcare, and policy decisions. Although the need for purposeful replication of systematic reviews has been noted previously,<sup>31 32</sup> we provide here a structured approach to support this. Our systematic review replication checklist aims to provide guidance to authors, commissioners, funders, and other users of systematic reviews, including clinicians, patients, and representatives from organisations involved with policy making. Systematic reviews can often be subject to errors and ambiguity or uncertainty, owing to inaccurate data collection or analysis, as well as judgments and bias in how data are collected, analysed, and interpreted (eg, breadth of the question, criteria for inclusion, variables for adjustment, decisions about how data are combined). Replication should be optimised for appropriate selection of systematic reviews to replicate, while eliminating excessive waste of resources by unnecessary duplication.<sup>6 9</sup>

Because the checklist contains guidance considerations rather than formal mandatory items with only one correct approach, users must consider the relevance of each item, from the perspective of appropriate stakeholder. Use of the checklist might also stimulate additional questions that look at the unique characteristics related to replication in the substantive domain of interest. Users' increased familiarity with other supporting tools (eg, for prioritisation or quality assessment) will also facilitate use of this checklist. Most importantly, making a decision on whether to replicate requires a good understanding of the underlying body of evidence—that is, the extent, quantity, and quality of existing review (and primary study) results. This approach could require preliminary work, such as a scoping review.

The purpose of replication is to reduce but not necessarily to eliminate decisional uncertainty for translation of results into policy or practice. Identifying the specific sources of uncertainty to be resolved by the replication is therefore a key component of the checklist and will drive the design of the replication. Depending on the specific uncertainty or controversy, repeating only part of the review might be warranted. For example, when critics raised concerns about a systematic review on vaccines for human papillomavirus owing to the exclusion of unpublished reports,<sup>33</sup> the authors conducted a replication that focused mainly on a reanalysis after adding previously omitted data.<sup>34</sup> When author influence is a source of uncertainty, it is imperative to have an independent team conduct the replication. In other instances, author overlap between the original and replication reviews might be acceptable. All these associations should be declared in the published review. When considering each checklist item, users need to articulate what concerns or uncertainties of the previous reviews should be assessed and how a replication would resolve these concerns. Designing

replications to resolve specific sources of uncertainty in this way will limit the waste of research resources while contributing directly to the advancement of knowledge.

Recent guidance on when to consider updating systematic reviews<sup>13</sup> is relevant for assessing when replication of previous systematic reviews is likely to have value, for example, when new evidence has become available. However, such guidance does not ensure that uncertainties around the many judgments or choice of technical methods are addressed.

### Strengths

This checklist can help researchers decide whether scarce research resources should be dedicated to purposefully replicating previous systematic reviews to resolve uncertainties in the existing body of knowledge. The checklist was based on the input from multiple stakeholders including methodological experts, following a literature review, interviews, a structured survey, and a consensus meeting. A rigorous process was followed in consultation with representatives from key related disciplines and perspectives.

A varied set of examples of useful replication and wasteful duplication were used to ensure that the checklist is pragmatic. The Campbell Collaboration, Cochrane, and others represented at the meeting are now proposing to use the checklist to guide their decisions on when replication is warranted.

### Limitations

The usability, acceptability, and usefulness of the checklist will need to be assessed and tailored to different users. For example, we have proposed supporting tools, but others could be more suitable for specific review questions or users. Use of the checklist for other types of review questions beyond interventions needs to be evaluated. Furthermore, valid responses to the items depend on input from experts in both systematic review methods, as well as in the content area of the review question; both sources of information could be biased.

Application of the four item checklist requires time and human resources. However, our large team of stakeholders thought that the investment of resources for completing the checklist would be well justified given the potential value of replicating systematic reviews that inform policy and practice as well as the opportunity to limit waste related to unnecessary duplication.

Finally, the use of this checklist depends on stakeholders recognising the value of systematic review replication. Promoting the value of systematic review replication will be an important element in the dissemination and implementation of the checklist.

### Conclusions and next steps

This work emphasises the importance of systematic review replication as a necessary and useful element in the advancement of knowledge, along with the use of updates, living systematic reviews, and open evidence

synthesis. In view of the role of systematic reviews in policy making and guideline development, the validity and reliability of these updated findings should be tested. The checklist serves as an explicit prompt to carefully consider the value of replication alongside other options such as updating, new reviews, and overviews of reviews.

We will test and tailor the usability, acceptability, or usefulness of the tool with specific user groups as needed. A more detailed framework for identifying uncertainties and designing the replication response<sup>35</sup> is also being proposed as part of the guidance on how to replicate systematic reviews. Since this work focused on systematic reviews of interventions, its application to other types of review questions and involving other types of systematic reviews (eg, qualitative, scoping, diagnosis, and prognosis) should be explored.

### Author affiliations

<sup>1</sup>Department of Medicine, University of Ottawa, 501 Smyth Road, Room L1227, Ottawa, ON, K1H 8L6, Canada

<sup>2</sup>Bruyere Research Institute, Ottawa, ON, Canada

<sup>3</sup>Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada

<sup>4</sup>School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada

<sup>5</sup>Department of Internal Medicine, American University of Beirut, Beirut, Lebanon

<sup>6</sup>Public Health Agency of Canada, Ottawa, ON, Canada

<sup>7</sup>Hospital for Sick Children, Toronto, ON, Canada

<sup>8</sup>Department of Medicine, University of Toronto, Toronto, ON, Canada

<sup>9</sup>BMJ Editorial, London, UK

<sup>10</sup>Pan American Health Organization (PAHO/WHO), Unit of Health Services and Access, Washington, DC, USA

<sup>11</sup>Faculty of Health, Dalhousie University, Halifax, NS, Canada

<sup>12</sup>Stanford Prevention Research Center, Department of Medicine and Department of Epidemiology and Population Health, Stanford University, Stanford, CA, USA

<sup>13</sup>JBI, Faculty of Health and Medical Sciences, University of Adelaide, South Australia

<sup>14</sup>School of Rehabilitation Therapy, Queen's University, Kingston, ON, Canada

<sup>15</sup>Centre for Research in Educational and Community Services, School of Psychology, Faculty of Social Sciences, Ottawa, ON, Canada

<sup>16</sup>World Health Organization, Partnership for Maternal, Newborn, and Child Health (PMNCH), Geneva, Switzerland

<sup>17</sup>Ingersoll, ON, Canada

<sup>18</sup>Department of Anesthesia and Perioperative Medicine, and Department of Epidemiology and Biostatistics, Western University, London, ON, Canada

<sup>19</sup>Department of Research in Biomedicine and Health, University of Split School of Medicine, Split, Croatia

<sup>20</sup>Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, ON, Canada

<sup>21</sup>NHMRC Clinical Trials Centre, University of Sydney, Sydney, NSW, Australia

<sup>22</sup>Faculty of Health, University of Plymouth, UK

<sup>23</sup>School of Public Health and Preventive Medicine, Monash University, Melbourne, VIC, Australia

<sup>24</sup>Cochrane Musculoskeletal Group, Ottawa Hospital Research Institute, Ottawa, ON, Canada

<sup>25</sup>Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, London, UK

<sup>26</sup>College of Education and Human Development, Georgia State University, Atlanta, GA, USA

- <sup>27</sup>Department of Family Medicine, University of Ottawa, Ottawa, ON, Canada
- <sup>28</sup>Epistemonikos Foundation, Santiago, Chile
- <sup>29</sup>Canadian Agency for Drugs and Technologies in Health, Ottawa, ON, Canada
- <sup>30</sup>Narendra Paul Loomba Department of Management, Baruch College, New York, NY, USA
- <sup>31</sup>Cochrane Canada and McMaster GRADE Centres, Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada
- <sup>32</sup>Population Health, Interdisciplinary School of Health Sciences, University of Ottawa, Ottawa, ON, Canada
- <sup>33</sup>Department of Cardiovascular Medicine, Mayo Clinic, Rochester, MN, USA
- <sup>34</sup>Ottawa, ON, Canada
- <sup>35</sup>Cochrane, London, UK
- <sup>36</sup>School of Rehabilitation, Faculty of Medicine, Université de Montréal, Montreal, Canada
- <sup>37</sup>University of Louisville, Louisville, KY, USA
- <sup>38</sup>Department of Mechanical Engineering, University of Ottawa, Ottawa, ON, Canada
- <sup>39</sup>Cardiovascular Research Methods Centre, University of Ottawa Heart Institute, Ottawa, ON, Canada
- <sup>40</sup>George Mason University, Fairfax, VA, USA
- <sup>41</sup>Campbell Collaboration, Oslo, Norway
- We thank Paul Glasziou, Simon Lewin, and Tammy Clifford for their comments on this work; and Jiakuan Wang and Humaira Mahfuz for collecting and summarising online survey data.
- Contributors:** PT and VAW contributed equally to this paper. PT, VAW, and HW conceptualised the research question. SK, VAW, PT, LJM, JP, and JPP operationalised the study design and collected the data. SK, VAW, and PT drafted the manuscript. All authors provided input on the direction of the study, content of the checklist, and two rounds of manuscript revision. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.
- Funding:** This work was funded by an operating grant from the Canadian Institutes of Health (PJT-148870). The funders had no role in the study design, the collection, analysis, and interpretation of data or the writing of the article or the decision to submit it for publication.
- Competing interests:** All authors have completed the ICMJE uniform disclosure form at [www.icmje.org/doi\\_disclosure.pdf](http://www.icmje.org/doi_disclosure.pdf) and declare: support from the Canadian Institutes of Health Research for the submitted work. PT holds a Canada Research Chair in health equity. LGC works for the Pan American Health Organization/World Health Organization, an intergovernmental public health organisation part of the United Nations system; contributions to this publication were done in his personal time and do not necessarily reflect decisions or policies of his employer. BH reports honorariums from Cornerstone Research Group, outside the submitted work, for methodological advice related to the conduct of systematic reviews and meta-analysis. JPP is a member of the Governing Board of Cochrane, a producer of systematic reviews. KSW is full time employee of Cochrane. In addition, all coauthors have a direct or indirect interest in systematic reviews and replication as part of their job or academic career.
- Data sharing:** Background material for the consensus meeting are available on request from the corresponding author.
- PT affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.
- Peer review and provenance:** Not commissioned; externally peer reviewed.
- Collins FS, Tabak LA. Policy: NIH plans to enhance reproducibility. *Nature* 2014;505:612-3. doi:10.1038/505612a
  - Alberts B, Cicerone RJ, Fienberg SE, et al. SCIENTIFIC INTEGRITY. Self-correction in science at work. *Science* 2015;348:1420-2. doi:10.1126/science.aab3847
  - Pan American Health Organization. 49th Directing Council. 61st Session of the Regional Committee of WHO for the Americas. Policy on research for health PAHO. Washington, DC; 2009. Report No: document CD49/10.
  - Wanous JP, Sullivan SE, Malinak J. The role of judgment calls in meta-analysis. *J Appl Psychol* 1989;74:259-64. doi:10.1037/0021-9010.74.2.259
  - Doundoulakis I, Antza C, Apostolidou-Kiouti F, et al. Overview of systematic reviews of non-vitamin K oral anticoagulants in atrial fibrillation. *Circ Cardiovasc Qual Outcomes* 2018;11:e004769. doi:10.1161/CIRCOUTCOMES.118.004769
  - Siontis KC, Ioannidis JPA. Replication, duplication, and waste in a quarter million systematic reviews and meta-analyses. *Circ Cardiovasc Qual Outcomes* 2018;11:e005212. doi:10.1161/CIRCOUTCOMES.118.005212
  - Weir MC, Grimshaw JM, Mayhew A, Fergusson D. Decisions about lumping vs. splitting of the scope of systematic reviews of complex interventions are not well justified: a case study in systematic reviews of health care professional reminders. *J Clin Epidemiol* 2012;65:756-63. doi:10.1016/j.jclinepi.2011.12.012
  - Worswick J, Wayne SC, Bennett R, et al. Improving quality of care for persons with diabetes: an overview of systematic reviews - what does the evidence tell us? *Syst Rev* 2013;2:26. doi:10.1186/2046-4053-2-26
  - Ioannidis JPA. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Milbank Q* 2016;94:485-514. doi:10.1111/1468-0009.12210
  - Page MJ, Moher D. Mass production of systematic reviews and meta-analyses: An exercise in mega-silliness? *Milbank Q* 2016;94:515-9. doi:10.1111/1468-0009.12211
  - Siontis KC, Hernandez-Boussard T, Ioannidis JPA. Overlapping meta-analyses on the same topic: survey of published studies. *BMJ* 2013;347:f4501. doi:10.1136/bmj.f4501
  - Moher D, Tsertsvadze A. Systematic reviews: when is an update an update? *Lancet* 2006;367:881-3. doi:10.1016/S0140-6736(06)68358-X
  - Garner P, Hopewell S, Chandler J, et al. Panel for updating guidance for systematic reviews (PUGs). When and how to update systematic reviews: consensus and checklist. *BMJ* 2016;354:i3507. doi:10.1136/bmj.i3507
  - Schünemann HJ, Wiercioch W, Brozek J, et al. GRADE Evidence to Decision (EtD) frameworks for adoption, adaptation, and de novo development of trustworthy recommendations: GRADE-ADOLOPMENT. *J Clin Epidemiol* 2017;81:101-10. doi:10.1016/j.jclinepi.2016.09.009
  - Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. *PLoS Med* 2010;7:e1000217. doi:10.1371/journal.pmed.1000217
  - Jull J, Graham ID, Kristjansson E, et al. Members of the CONSORT-Equity 2017 and Boston Equity Symposium. Taking an integrated knowledge translation approach in research to develop the CONSORT-Equity 2017 reporting guideline: an observational study. *BMJ Open* 2019;9:e026866. doi:10.1136/bmjopen-2018-026866
  - Minelli C, Baio G. Value of information: a tool to improve research prioritization and reduce waste. *PLoS Med* 2015;12:e1001882. doi:10.1371/journal.pmed.1001882
  - Hoomans T, Seidenfeld J, Basu A, Meltzer D. Systematizing the Use of Value of Information Analysis in Prioritizing Systematic Reviews. Rockville (MD): Agency for Healthcare Research and Quality (US); 2012. (AHRQ Methods for Effective Health Care). <https://www.ncbi.nlm.nih.gov/books/NBK107294/>
  - 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. doi:10.1136/bmj.j4008
  - Akl EA, Fadlallah R, Ghandour L, et al. The SPARK Tool to prioritise questions for systematic reviews in health policy and systems research: development and initial validation. *Health Res Policy Syst* 2017;15:77. doi:10.1186/s12961-017-0242-4
  - Cowan K, Oliver S. JLA Guidebook | James Lind Alliance. version 5. Southampton: National Institutes for Health Research Evaluation, Trials and Studies Coordinating Centre; 2016:95. <https://www.jla.nihr.ac.uk/jla-guidebook/>
  - Rudan I, Gibson JL, Ameratunga S, et al. Child Health and Nutrition Research Initiative. Setting priorities in global child health research investments: guidelines for implementation of CHNRI method. *Croat Med J* 2008;49:720-33. doi:10.3325/cmj.2008.49.720
  - Whiting P, Savović J, Higgins JPT, et al. ROBIS group. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol* 2016;69:225-34. doi:10.1016/j.jclinepi.2015.06.005
  - Hall A, Hewitt G, Tuffrey V, de Silva N. A review and meta-analysis of the impact of intestinal worms on child growth and nutrition. *Matern Child Nutr* 2008;4(Suppl 1):118-236. doi:10.1111/j.1740-8709.2007.00127.x
  - Tovey D, Littell JH, Grimshaw JM. The end of the wormwars? *Cochrane Database Syst Rev* 2016;9:ED000116.
  - Aiken AM, Davey C, Hargreaves JR, Hayes RJ. Re-analysis of health and educational impacts of a school-based deworming programme

- in western Kenya: a pure replication. *Int J Epidemiol* 2015;44:1572-80. doi:10.1093/ije/dyv127
- 27 Taylor-Robinson DC, Maayan N, Soares-Weiser K, Donegan S, Garner P. Deworming drugs for soil-transmitted intestinal worms in children: effects on nutritional indicators, haemoglobin, and school performance. *Cochrane Database Syst Rev* 2015;(7):CD000371. doi:10.1002/14651858.CD000371.pub6
  - 28 Augustincic Polec L, Petkovic J, Welch V, et al. Strategies to increase the ownership and use of insecticide-treated bednets to prevent malaria. *Cochrane Database Syst Rev* 2015;(3):CD009186. doi:10.1002/14651858.CD009186.pub2
  - 29 Jadad AR, Cook DJ, Browman GP. A guide to interpreting discordant systematic reviews. *CMAJ* 1997;156:1411-6.
  - 30 Moja L, Fernandez del Rio MP, Banzi R, et al. Multiple systematic reviews: methods for assessing discordances of results. *Intern Emerg Med* 2012;7:563-8. doi:10.1007/s11739-012-0846-1
  - 31 Scott AM, Beller E, Glasziou P, et al. Is antimicrobial administration to food animals a direct threat to human health? A rapid systematic review. *Int J Antimicrob Agents* 2018;52:316-23. doi:10.1016/j.ijantimicag.2018.04.005
  - 32 Low J, Ross JS, Ritchie JD, et al. Comparison of two independent systematic reviews of trials of recombinant human bone morphogenetic protein-2 (rhBMP-2): the Yale Open Data Access Medtronic Project. *Syst Rev* 2017;6:28. doi:10.1186/s13643-017-0422-x
  - 33 Jørgensen L, Gøtzsche PC, Jefferson T. The Cochrane HPV vaccine review was incomplete and ignored important evidence of bias[correction in: *BMJ Evid Based Med* 2020;25:e3]. *BMJ Evid Based Med* 2018;23:165-8. doi:10.1136/bmjebm-2018-111012
  - 34 Tovey D, Soares-Weiser K. Cochrane's Editor in Chief responds to a BMJ Evidence-Based Medicine article criticizing the Cochrane Review of HPV vaccines. 2018. [https://www.cochrane.org/sites/default/files/public/uploads/cochrane\\_hpv\\_response\\_3sep18.pdf](https://www.cochrane.org/sites/default/files/public/uploads/cochrane_hpv_response_3sep18.pdf)
  - 35 Grimshaw JM, White H, Piggott T, Kristjansson E, Tugwell P, Welch VA. Developing guidance for when and how to replicate systematic reviews. What works. Global Summit, Mexico City. 16-18 October, 2019.

**Web appendix:** Supplementary appendix