

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	Economic evaluations of scaling up strategies of evidence-based HEALTH interventions: a systematic review Protocol
<b>AUTHORS</b>	Brundisini, Francesca; Zomahoun, Hervé Tchala Vignon; Légaré, France; Rhéault, Nathalie; Bernard-Uwizeye, Claude; Massougbojji, José; Gogovor, Amédé; Tchoubi, Sébastien; Assan, Odilon; Laberge, Maude

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Dams, Judith University Medical Center Hamburg-Eppendorf
<b>REVIEW RETURNED</b>	07-May-2021

<b>GENERAL COMMENTS</b>	<p>The authors aim to perform a systematic review on scaling up strategies of evidence-based health interventions to describe the state of art of scaling up strategies used in economic evaluations as well as to identify trade-offs of scaling up strategies. This topic is particularly important as it addresses the transition between scientific results and their implementation in practice. Thus, future researchers but also decision makers on economic issues could benefit from the results.</p> <p>The review's broad scope is remarkable: it is not restricted by publication date, country, intervention or type of study. Thus, it is to be expected that the results will be transferable to many areas. The proposed methodology also follows the high standards of evidence-based medicine. The expertise in this can be seen in the structure of the protocol alone, starting with the inclusion criteria based on PICOS, through the detailed presentation of the search and data extraction, to the quality assessment.</p> <p>I would therefore only address one minor point: A large heterogeneity in the results is expected (as already mentioned by the authors). How should this be dealt with? Is it nevertheless to be expected that generally valid recommendations for action can be derived for future research or decision-makers?</p>
-------------------------	--

<b>REVIEWER</b>	Jacobsen, Elisabet University of Aberdeen
<b>REVIEW RETURNED</b>	04-Jun-2021

<b>GENERAL COMMENTS</b>	<p>Thank you for this well conducted protocol which aims to highlight (methodological) lessons learnt from existing economic evaluations of broadening access to effective health interventions which will hopefully benefit future economic evaluations.</p> <p>I have a few comments on the protocol:</p>
-------------------------	---

	<p>Abstract: Please consider revising the first sentence. Firstly, it is mentioned that this is a systematic review of characteristics and methods applied in economic evaluations. Because study characteristics and methods applied are data extraction items that will be obtained from these economic evaluations (which is mentioned later on) this could be left out from this sentence. Lastly, please consider rephrasing “scaling up science” to clarify what you mean.</p> <p>Page 7: Please elaborate on some of the complexities of implementing scaling up strategies that you mention here. E.g. what additional costs are occurred? Are all these complexities unique to “scaling up” strategies? Some of the study heterogeneities mentioned here are true for most economic evaluations. It would benefit this paragraph if the focus was on the study heterogeneities that are unique to economic evaluations assessing the expansion strategies.</p> <p>Data collection process: This protocol would benefit from the addition of the data extraction template that will be used for the pilot test.</p> <p>The Drummond and Jefferson checklist is usually used for assessing the methodological quality of economic evaluations conducted alongside single effectiveness studies. Please clarify whether the modelling studies will also be quality assessed, and if so, how they will be quality assessed.</p> <p>Search strategy: How all-encompassing is “scaling up” as a search term for the type of strategies that you are focusing on in this systematic review of economic evaluations. My understanding is that this is a systematic review of economic evaluations on strategies expanding the availability of already established effective health interventions. The protocol would benefit from a brief explanation of why “scaling up” as a search term (or any of the other scaling keywords) will identify these interventions or how these keywords were identified by the team.</p>
--	---

## VERSION 1 – AUTHOR RESPONSE

### Reviewers’ Comments to Authors:

#### Reviewer 1

A large heterogeneity in the results is expected (as already mentioned by the authors). How should this be dealt with?

Authors’ Response: Thank you for pointing this out. We have addressed this aspect in the data synthesis section (page 15). The added text reads as follows:

*“We will explore this heterogeneity by narratively synthesizing the differences, and if possible, the similarities in settings, participants, intervention, comparison and outcome characteristics across included studies. For example, we will perform the data synthesis of economic evaluation methods according to the economic evaluation parameters reported”*

Is it nevertheless to be expected that generally valid recommendations for action can be derived for future research or decision-makers?

Authors' Response: Yes, we believe that generally valid recommendations for action can be derived for future research and/or decision-makers. We believe that we have already pointed to potential actionable recommendations in the main text (page 16), but we revised the wording to clarify these recommendations. Revised text (in blue) reads as follows:

*As such, we believe that the findings of this study will point to identify valid recommendations for action for future research and decision-makers. First, this study can help guide future research aimed at defining costing tools and models that can be easily used in scaling up frameworks and plans. It will contribute to define the nature and selection of costs that are integral to the successful roll out of EBIs on large scale, as well as the benefits and disadvantages of each economic methodological approaches aimed at evaluating strategies identified in the literature. Second, as scaling science is becoming an increasingly relevant area for research, policy, and practice, clarifying how underlying methodological assumptions are based on evidence and on the multi-factorial complexity of real-world scaling strategies, thus, will advance the quantity and quality of the information extractable from the evidence to inform both research and practice. We believe this review will then offer opportunities for improvement in the quality, production, reporting, and application in practice of health economic evaluative methods to scaling up strategies.*

## Reviewer 2

Abstract: Please consider revising the first sentence. Firstly, it is mentioned that this is a systematic review of characteristics and methods applied in economic evaluations. Because study characteristics and methods applied are data extraction items that will be obtained from these economic evaluations (which is mentioned later on) this could be left out from this sentence. Lastly, please consider rephrasing “scaling up science” to clarify what you mean.

Authors' Response: We think this is an excellent suggestion. First, we revised the wording for “scaling up science” to simply “scaling science” as indicated in the literature, see (Simmons, Fajans et al. 2007, Milat, King et al. 2011, McLean and Gargani 2019). Second, we revised the first sentence and removed the part as suggested by the reviewer. The revised text reads as follows (page 3):

*Scaling science aims to help roll out evidence-based research results on a wide scale to benefit more individuals. Yet, little is known on how to evaluate economic aspects of scaling up strategies of evidence-based health interventions.*

Page 7: Please elaborate on some of the complexities of implementing scaling up strategies that you mention here. E.g. what additional costs are occurred? Are all these complexities unique to “scaling up” strategies? Some of the study heterogeneities mentioned here are true for most economic evaluations. It would benefit this paragraph if the focus was on the study heterogeneities that are unique to economic evaluations assessing the expansion strategies.

Authors' Response: We appreciate the reviewer's feedback. We have now added further details regarding the lack of clarity and standardized conceptualizations of economic constructs in implementation and scaling up frameworks.

Section not modified:

*Oftentimes, the lack of complete availability of scaling up cost data or the use of models leads economic analysts to rely on assumptions that may not reflect the complexity of implementing scaling up strategies. For example, economic evaluations may posit that scaling-up implementation costs are a fixed part of the*

*intervention costs. In reality, scaling up strategies may present additional costs to that of the intervention that can greatly vary across interventions and settings, potentially leading to both economies and/or diseconomies of scale.<sup>29</sup> Costs and cost-effectiveness estimates may change according to the type of intervention being expanded, the size of the targeted population, the prevalence/incidence of the disease, the relevant efficacy level of the intervention, the geography, and the financial resources available and needed. Specific to scaling up strategies, costs and estimates related to infrastructure and available human resources can vary based on the different scaling up strategy operationalization and management, the cost impacts of change, including the excess cost of service delivery as uptake changes and the opportunity costs to providers and patients participating in the activities.*

Additional text to add clarity:

*Finally, implementation and scale-up theoretical frameworks – that support thinking and interpretation of “real world” complex data – consider economic constructs in scaling up strategies in different ways. For example, some frameworks consider cost (and resource) mobilisation as a key objective,<sup>33 34</sup> yet implementation frameworks consider costs as an implementation outcome.<sup>35</sup> Frameworks vary also in the ways they consider potential benefit or effectiveness (‘Cost-benefit’).<sup>36</sup>*

Data collection process: This protocol would benefit from the addition of the data extraction template that will be used for the pilot test.

Authors’ Response: We think this is an excellent suggestion. We have added the data extraction template as a supplemental file (3) and reported it in the text at page 13.

The Drummond and Jefferson checklist is usually used for assessing the methodological quality of economic evaluations conducted alongside single effectiveness studies. Please clarify whether the modelling studies will also be quality assessed, and if so, how they will be quality assessed.

Authors’ Response: We are grateful for this comment as it points to an important issue in the field. Methodological studies on economic evaluation checklists extensively argue how there is no single checklist deemed to be complete or fit for all types of economic evaluation studies. In the latest years, experts have tried to make these checklists universal, but many would argue that we are not there yet, as many tend to customize the list according to the type of economic evaluations considered in their study.(Gerkens, Crott et al. 2008, Walker, Wilson et al. 2012, Watts and Li 2019)

The Drummond and Jefferson checklist (a 35-item checklist) is one of the most commonly used checklists,(Gerkens, Crott et al. 2008, Walker, Wilson et al. 2012, Watts and Li 2019). We selected this checklist because it provides a detailed breakdown of appraisal items and allows to consider also different types of economic evaluations (both full and partial, modelling studies as well as economic commentaries, such as methodological studies).(Gerkens, Crott et al. 2008, Walker, Wilson et al. 2012, Watts and Li 2019) Considering the broad scope of this systematic review and the inclusion of a wide range of economic evaluation approaches, then, this tools appears to be the most comprehensive and best suited to appraise them.

We came to this decision after reviewing other checklists used in the field. Among the most used and validated checklists, we excluded the CHEERS tool because it is a reporting assessment tool and not a methodological assessment tool (it was created with the purpose of assessing the way the economic evaluations are reported),(Husereau, Drummond et al. 2013) the CHEC tool was specifically designed only for trial-based and observational studies with no items for other studies, especially modelling studies (all the authors listed above warn to proceed with caution when using this tool with modelling studies, suggesting not to use it).(Evers, Goossens et al. 2005) The Phillips tool is

a 65-item list only for modelling studies. It doesn't allow us to appraise the quality of other types of economic evaluation studies. (Philips, Bojke et al. 2006) Eventually, this might be considered if the majority of included studies will be modelling studies, which at this point in time we don't know yet. The QHES checklist might be a good alternative as it allows to include different types of studies, and it is quite simple. However, methodological papers addressing and comparing various economic evaluation assessment tools highlight some issues with this tool. (Chiou, Hay et al. 2003) In particular, the checklist is quite simple and groups several quality criteria together, limiting the opportunity to determine a score that actually reflects the quality among all the studies (because it has a yes/no scoring system). If our systematic review was narrower in scope and retrieved similar studies, then we might have considered this checklist. Finally, since there aren't many economic evaluations of scaling up strategies, nor of implementation studies (the literature is way smaller compared to traditional health technology economic evaluations), there is no clear precedence on which economic evaluation appraisal tool is the best in this field.

Search strategy: How all-encompassing is "scaling up" as a search term for the type of strategies that you are focusing on in this systematic review of economic evaluations. My understanding is that this is a systematic review of economic evaluations on strategies expanding the availability of already established effective health interventions. The protocol would benefit from a brief explanation of why "scaling up" as a search term (or any of the other scaling keywords) will identify these interventions or how these keywords were identified by the team.

Authors' Response:

Thank you for this comment. In fact, an inaccuracy crept into our protocol. The exact concepts used for the review are : 1) Scaling (and not scaling **up**) and 2) Economic Evaluation. The concept Scaling includes expressions like "scaling up", "scaling out" or "scale up". In using the adjacency, we were also able to combine words that have the meaning "Spread" with words that have the meaning "Innovation". This operation enables to have countless of variants like, for example, "spread of technologies", "widespread adoption of the technology", "rolling out the model of care". The concept related to health interventions was not used in the search strategy in order to be inclusive and to identify them during the study selection process. For more clarity, we will add the search strategy in the supplementary material.

We also modified the text at page 11 for describe the process used for select the terms used in the search strategy:

*The search will include a combination of the following two concepts: 1) scaling and 3) Economic Evaluation. No language restrictions will be applied. The search strategy in Ovid Medline is in the Supplementary Materials.*

*The following source were used to find the search terms: 1) Previous reviews who used the concept of scaling up and the concept of economic evaluation; 2) Search strategies created for two other reviews at the Unité de Soutien SRAP and who are not published at yet 3) The knowledge of the experts of the team in scaling up 4) The thesaurus of the consulted bibliographic databases. All words and expressions found were tested and evaluated by the information specialist before to be integrated or rejected in the search strategy. The search strategy was commented by the others members of the team, and a final version was produced afterwards.*

*The concept Scaling was created for retrieved all the potential expressions for designed the idea of the spreading of an innovation. It is designed to retrieved very used expression like "scaling up", "scale up", "spread of technologies", but also many variations like "widespread adoption of the*

*technology" or "rolling out the model of care". The concept of Economic Evaluation integrated all synonyms like "cost evaluation", "economic analysis" and "net benefit".*

Cited references :

- Chiou, C. F., J. W. Hay, J. F. Wallace, B. S. Bloom, P. J. Neumann, S. D. Sullivan, H. T. Yu, E. B. Keeler, J. M. Henning and J. J. Ofman (2003). "Development and validation of a grading system for the quality of cost-effectiveness studies." Med Care **41**(1): 32-44.
- Evers, S., M. Goossens, H. de Vet, M. van Tulder and A. Ament (2005). "Criteria list for assessment of methodological quality of economic evaluations: Consensus on Health Economic Criteria." Int J Technol Assess Health Care **21**(2): 240-245.
- Gerkens, S., R. Crott, I. Cleemput, J. P. Thissen, M. C. Closon, Y. Horsmans and C. Beguin (2008). "Comparison of three instruments assessing the quality of economic evaluations: a practical exercise on economic evaluations of the surgical treatment of obesity." Int J Technol Assess Health Care **24**(3): 318-325.
- Husereau, D., M. Drummond, S. Petrou, C. Carswell, D. Moher, D. Greenberg, F. Augustovski, A. H. Briggs, J. Mauskopf and E. Loder (2013). "Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement." BMJ **346**(mar25 1): f1049-f1049.
- McLean, R. and J. Gargani (2019). Scaling impact: innovation for the public good, Routledge.
- Milat, A. J., L. King, A. Bauman and S. Redman (2011). "Scaling up health promotion interventions: an emerging concept in implementation science." Health Promot J Austr **22**(3): 238.
- Philips, Z., L. Bojke, M. Sculpher, K. Claxton and S. Golder (2006). "Good Practice Guidelines for Decision-Analytic Modelling in Health Technology Assessment." PharmacoEconomics **24**(4): 355-371.
- Simmons, R., P. Fajans and L. Ghiron (2007). Scaling up health service delivery: from pilot innovations to policies and programmes. Geneva: World Health Organization.
- Walker, D. G., R. F. Wilson, R. Sharma, J. Bridges, L. Niessen, E. B. Bass and K. Frick (2012). AHRQ Methods for Effective Health Care. Best Practices for Conducting Economic Evaluations in Health Care: A Systematic Review of Quality Assessment Tools. Rockville (MD), Agency for Healthcare Research and Quality (US).
- Watts, R. D. and I. W. Li (2019). "Use of Checklists in Reviews of Health Economic Evaluations, 2010 to 2018." Value Health **22**(3): 377-382.