Author's response to reviews

Title: Policymakers' and other stakeholders' perceptions of key considerations for health system decisions and the presentation of evidence to inform those considerations: an international survey

Authors:

Joshua P Vogel (vogeljo@who.int)
Andrew D Oxman (oxman@online.no)
Claire Glenton (claire.glenton@kunnskapssenteret.no)
Sarah Rosenbaum (sarah@rosenbaum.no)
Simon Lewin (simon.lewin@kunnskapssenteret.no)
A. Metin Gülmezoglu (gulmezoglu@who.int)
João Paulo Souza (souzaj@who.int)

Version: 2 Date: 4 May 2013

Author's response to reviews: see over
Dear Sir/Madam,

Many thanks for your consideration and the comments from your reviewers on our manuscript:

“Policymakers’ and other stakeholders’ perceptions of key considerations for health system decisions and the presentation of evidence to inform those considerations: an international survey”

We have addressed the issues raised by your reviewers and we trust you will find the article improved as a result.

Response to Reviewers

Reviewer 1

1.1. It would be useful to elaborate on the sampling framework that was used to select participants in each country. In terms of the 10 participants target in each country, was the intent to have a mix of all stakeholders in each country. It seems that in some countries, some stakeholder groups did not respond. Or is it that partners have more established links with some stakeholder groups in some countries, thus affected the overall response.

We recruited participants for diversity, rather than representativeness, using the network of countries participating in DECIDE and SURE. However, participants in the survey were not involved or familiar with the DECIDE project. We have elaborated on the sampling framework in the methods section, paragraph 1:
We conducted an international online survey to determine perceptions of the importance and use of the criteria within the DECIDE framework and of evidence summaries. We aimed to survey a diverse (rather than representative) group of policymakers from different countries, with a wide range of experience with different types of health policy and management decisions and with different perspectives.

1.2 There was no indication of whether the survey was pilot tested or translated to another language (was that needed given that respondents were from several countries)

The survey was pilot tested with a small group of participants and subsequently reviewed; the survey was conducted in English only. We have adjusted the last few sentences of paragraph one in the Methods section to reflect this:

We prepared an online survey (in English only) using the LimeService online platform [https://www.limeservice.com/] (see Appendix for survey questions). The survey was revised after pilot testing in a small group of policymakers who provided feedback on the content, length, clarity and ease of use. Informed consent was obtained from survey participants prior to commencing the survey and results were de-identified when exported for analysis to protect confidentiality.

1.3 Given that the survey was distributed to diverse respondents, it would be useful to report findings by type respondent as this may impact their views and perception on some questions.

We agree with this observation. We had explored the data to look for variation by type of respondents, however no significant variation was evident. We agree it would be useful to demonstrate these findings. We have modified the final paragraph of the Methods section:

Our primary analysis focused on implications for our evidence to health system decisions framework and evidence summaries. We explored potential differences in responses across participants from different countries (DECIDE versus SURE partner countries) and across groups with different types of experience (with versus without research training or experience).

We have added Supplementary Table 1 to show these findings, and added the following statement to the final paragraph in the Results section:

Overall, there was relatively little variation in the responses. Nonetheless, we explored potential differences in responses across participants from different countries and across groups with different types of research experience (Supplementary Table 1). We did not find any apparent differences in responses based on the respondents’ country (DECIDE versus SURE partner countries) or experience (with versus without research training or experience).

1.4 The survey was administered online and as such informed consent would have been challenging. What measures did the authors take to maintain respondent confidentiality? The limesurvey application includes a page that shows which respondents completed the survey which may have some ethical implications.
Informed consent was obtained from survey participants prior to commencing the survey and results were de-identified when exported for analysis to protect confidentiality. We have added a statement at the end of paragraph 1 of the Methods section to reflect this:

Informed consent was obtained from survey participants prior to commencing the survey and results were de-identified when exported for analysis to protect confidentiality.

1.5 Was the protocol for the study approved by an ethical review committee / board?

In accordance with the Norwegian Research Ethics Act of 30 June 2006 and the Act on Medical and Health Research (the Health Research Act) of 20 June 2008, the survey is outside the remit of the Act on Medical and Health Research and therefore could be implemented without approval from the Regional Committee for Medical Research Ethics.

1.6 How can authors explain about the partially completed surveys

We believe the length of the survey may have been a disincentive to completion, given that this group had registered and partially completed the survey. We added the following statement to paragraph 5 of the Discussion to reflect this:

The strengths of this survey were a good response rate from a diverse range of countries, backgrounds, levels of decision-making and organisations. The 25% of the surveys that were partially completed may have been due to survey length.

1.7 While around 90% of respondents stated that they knew what systematic review was, it would be useful to reflect on why only about 60% has used evidence from systematic reviews to inform decisions

We agree that highlighting this finding would be beneficial – it could be due to a lack of available evidence (despite a desire to use it) or other barriers to use of evidence in policy settings. We have added the following statement to the end of paragraph 2 of the Discussion to highlight this:

However, only 60.2% reported using systematic reviews to inform decision making. This may reflect both a lack of systematic reviews addressing relevant questions, as well as other reasons that have been found for not using research evidence to inform health policy decisions [18, 19].

1.8 Around 38% of respondents said that the grading systems should be the same for clinical and health system decisions. It would be good to mention about the characteristics of respondents. The same applies for those who indicated that different grading systems should be used

This point overlaps with the issue raised in point 1.3 above. We have added the following statement to the final paragraph of the Results section, as well as adding the findings by types of respondent to Supplementary Table 1.
We did not find any apparent differences in responses based on the respondents’ country (DECIDE versus SURE partner countries) or experience (with versus without research training or experience).

1.9 The conclusion can be strengthened by discussing about next steps for this work including user testing and the application of the framework to health system decisions. And if/how the results of this survey will inform user testing and further application of framework to decisions, hence the final selection of criteria included in the framework.

We agree with this point, and we have added the following paragraph at the end of the discussion:

This survey confirmed the relevance of the criteria that we had identified and incorporated in the DECIDE framework for health system decisions and suggests that the framework is likely to be helpful for informing health system decisions. Further development and evaluation of the framework will be based on practical applications of the framework to health system and population health decisions and user testing (1). Facets of the framework that will be addressed by user testing include its usefulness and usability (Box 1, adapted from (2) and (3)). Further work will address the advantages and disadvantages of using the same versus different systems for grading evidence, further clarification of the included criteria, the need for additional criteria, and the perceptions of policymakers and stakeholders who do not have a research or health professional background.

Box 1. Facets of the decision framework that will be addressed by user testing

<table>
<thead>
<tr>
<th>Facet</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Findability</td>
<td>can users locate what they are looking for?</td>
</tr>
<tr>
<td>2. Usefulness</td>
<td>does the framework have practical value for the user?</td>
</tr>
<tr>
<td>3. Usability</td>
<td>how easy and satisfying is the framework to use?</td>
</tr>
<tr>
<td>4. Understandability</td>
<td>do users understand the framework and the content correctly?</td>
</tr>
<tr>
<td>5. Credibility</td>
<td>is this framework/content trustworthy?</td>
</tr>
<tr>
<td>6. Desirability</td>
<td>is the framework something the user wants/responds positively to?</td>
</tr>
<tr>
<td>7. Identification</td>
<td>does the framework feel like it was designed for “someone like me (the user)”</td>
</tr>
</tbody>
</table>

Adapted from Morville et al (2) and Rosenbaum et al (3)

Reviewer 2.

2.1 From the perspective of understanding and ‘improving’ health system decision making, it could be argued that the paper contributes little in that it is hard to interpret the significance of the views of a diverse, but probably highly atypical group of policy ‘advisers’ (I suspect this would be a more accurate term for many of them rather than ‘policy makers’) who are also likely to have had significant prior exposure to the type of thinking represented by the authors of the survey.

We agree that the policymakers who responded were not representative of all policymakers and we have acknowledged this limitation in the discussion. The participants were selected to represent diverse views and experience. They represent a large range of countries,
qualifications, positions, and organizations. Their agreement regarding the DECIDE framework provides support for the relevance and comprehensiveness of the criteria incorporated in the framework and its potential usefulness. It is possible that policymakers with less research background might perceive the criteria and usefulness of the framework differently, but our findings do not support that conclusion (see response to comment 1.3 above) and we are unaware of other evidence that supports that conclusion. Work that is underway, including use of the framework by policymakers in a variety of organizations and user testing amongst different types of users will help to verify or refute the findings of this survey (see response to comment 1.9 above).

2.2 How are we to interpret the generally supportive views of the 112 respondents, the vast majority of whom had clinical backgrounds and research training/experience? It seems highly likely that those responding already had broadly positive views of the DECIDE framework and/or similar tools. It is further unlikely that such people would express much by way of dissent from the broad approach taken by DECIDE in a survey of this type. As a result, it is not clear how much further these findings take us.

As per point 1.1 above, survey participants had no exposure to the DECIDE framework prior to the survey nor did they have any other involvement in the DECIDE project. We were careful to frame all of the questions in the survey neutrally and to invite positive and negative responses. The assumption that respondents had broadly positive views of similar tools seems highly speculative, particularly given that neither we nor others have so far identified such a similar tool. As previously noted, the findings support the conclusion that the criteria we have identified are relevant to and appropriate for a wide range of health system decisions and that the framework is likely to be useful. They also support the conclusion that there is disagreement whether the same or different frameworks should be used to grade evidence.

2.3 The authors conclude that, ‘This survey supports the use of the DECIDE framework for health system decisions ...’, but can this claim be made on the basis of the responses to this type of survey? A more accurate formulation would be to say that the respondents to the survey expressed supportive attitudes to the use of the DECIDE framework.

We have changed this conclusion in the abstract to: Survey respondents were supportive of the DECIDE framework for health system decisions and the use of succinct summaries of the estimated size of effects and the quality of evidence.

2.4 We have little or no data on how the respondents act or intend to act in advising on or taking health system decisions in practice. We have no indication as to whether use of the DECIDE framework will alter decisions compared with other (perhaps more implicit) approaches, and we have no idea whether any altered decisions would confer benefits compared with some different (less rational) approach.

We agree. The survey did not address those questions, it does not answer them and we have made no claims in relation to those questions.

Reviewer 3.
3.1 The research question is interesting according to the increasing interest of evidence grading and recommendations. The objective of the study is relevant and warranted since the DECIDE framework is under development and should be evaluated before recommended and utilized in broader contexts. Other examples of evidence-grading and recommendations system could be presented, or at least mentioned, in the background.

*We have added the following to the third paragraph in the Background:*

[There are several different systems available to grade the quality of evidence about the effects of healthcare interventions.] Most of these have been used primarily for clinical practice guidelines and those systems have become increasingly similar to GRADE or replaced by GRADE (3-6). Other systems have been used for population health guidelines (primarily for public health rather than health system interventions), such as the systems used by the Task Force on Community Preventive Services (7-9), although the GRADE system is also used for public health and health system recommendations (10,11). All of these systems have focused primarily on grading the quality of evidence and only to a lesser extent, if at all, on frameworks for going from evidence to public health or health system decisions or recommendations.

3.2 The procedures are well described and the web-based survey is provided as an additional online file. However, the selection of participants is doubtful. The authors do not provide any information about the choice of the sampling frame. All countries invited to participate in the survey are working with DECIDE or SURE, and WHO is involved in the GRADE/DECIDE-system. Thus, a positive result could be expected.

*Language around selection of sampling frame has been refined in Methods paragraph 1, please see our response to comments 1.1 and 2.2 above.*

3.3 The response rate was 93% from the DECIDE partner countries. From WHO the response rate was 50% and from the SURE countries the response rate was 46%. Thus, the countries that are involved in the DECIDE-project are the ones that have given most responses in the survey, which gives an impression of bias. No differences in the responses from the different countries are described. The results in general are very positive. What would the result look like if the same questions had been asked to representatives from countries that are not familiar with evidence grading and recommendation systems? Were the two respondents that indicated that they disliked evidence grading systems from DECIDE countries?

*Clarified some of this in response to previous points. See response to comments 1.3 and 2.2 above. So far as we are aware, no countries are currently using the GRADE system for health system decisions or recommendations. As noted above in response to comment 3.1, GRADE is used by some groups for public health and health system recommendations, but the participants were not selected based on previous involvement in guideline development activities. Although it is likely that some participants had participated in guideline development activities and were familiar with GRADE, it is unlikely that many were, based on how they were identified and selected. Moreover, up to now the GRADE framework for going from evidence to recommendations has included only four criteria and there were no previous publications or information about the framework for going from evidence to health system.*
decisions or recommendations with the criteria that participants were asked about in this survey. The two respondents who disliked grading systems were from DECIDE countries.

3.4 The results shown in figure 1, 2 and 3 would be better presented in bar charts.

We agree with this observation, and have revised these three figures into cumulative bar charts.

3.5 The positive results are not unexpected due to potential study bias and should be questioned and discussed more thoroughly by the authors.

See responses to comments 1.1, 1.2, 2.2 and 3.1 above. As addressed in those responses, our aim was to select a diverse group of participants and we were careful to formulate the questions neutrally. We do not agree that the study was biased. We did not aim to obtain a representative sample and we have not claimed that the respondents were representative. We have discussed the limitations of the survey, including the response rate and the limited generalizability of the findings due to the respondents not being a representative sample.

3.6 Yes, but the aim of the study is not explicitly stated in the abstract.

Background of abstract revised to better reflect the objective:
The objective of this study was to determine policymakers’ perceptions regarding the criteria in the DECIDE framework and how best to summarise and present evidence to support health system decisions.

Many thanks for the opportunity to resubmit this article. We look forward to your forthcoming decision.

Best regards,

(signed)

Dr Joshua Vogel
vogeljo@who.int


