Reviewer's report

Title: Combining Evidence and Values using the Balance Sheet Method: the Effect of Deliberation on Priority Setting in a Low-income Country

Version: 1 Date: 23 July 2006

Reviewer: Richard Thomson

Reviewer's report:

General

This is an interesting study addressing the important and challenging issue of priority setting. Its focus on a low-income country adds an additional dimension of interest (and additional challenges). There is limited experience in this area (not only in low income countries) so publication of well applied and evaluated methods is sorely needed. This paper will be of potential value to both researchers and practitioners/policy makers in priority setting. The main thing that would make the paper more valuable would be greater clarity and detail on the why and how of this work, such that readers could more readily understand the approach and potentially duplicate or adapt for further research and/or policy applications.

Furthermore, it would seem that potentially rich material from recordings of the discussions and rankings was not utilised to understand the methods of decision making and the factors that influenced the rankings and to further explore the strengths and weaknesses of the approach - this is a missed opportunity to understand and interpret the findings. The paper would be much strengthened if the authors undertook some further structured qualitative analysis of the recordings of the sessions.

The emphasis on the difference between policy makers/professional and patient/public is an important element of this work, but again would benefit from further analysis/explanation.

-----------------------------------------------------------------------------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

More discussion of the search strategy would be valuable e.g. of the decision to limit the cost-effectiveness search to Tanzania or sub-Saharan Africa unless no records were identified. Could this have missed important data/studies? Would limiting to cost-effectiveness have missed important effectiveness studies? Similarly limiting to life years or DALYs? What was the source of the search strategy? Is limiting to Medline and Cochrane an issue? What methods were used to appraise study quality?

How was the literature on equity and political concerns undertaken? Search strategy? Yield? Methods of appraisal?

The reader (and anyone wishing to apply the approach) would benefit from more detail on the methods. In the â€œinterviewsâ€ were the data tables simply provided on paper or was there any form of presentation/discussion/explanation by researchers? There is a distinct possibility that the means of presentation of the evidence might affect the decision making. It would help if more detail were given about how this was done. What is meant by â€œPresentation of the evidence varied according to the informantâ€™s level of expertiseâ€?

The way that risks are presented can have a profound effect on interpretation and decision making with evidence that absolute risk is a better way of communicating than relative risk (as used in these tables). How clear was it that the participants understood the information they were given? Do the recordings or observation of the groupsâ€™ working help with this?

The actual use of the nominal group technique is not as clear as it might be. One expressed value of the nominal group technique is to get consensus in a way that allows all group members to have an equal say and in most cases â€œvotingâ€ would be done secretly. Why was voting only applied if a group couldnâ€™t reach consensus? And why was it open rather than secret voting? Donâ€™t these undermine the benefits of the nominal group method?
In this study it appears that the voting was open? What effect might this have on the results?

The sessions were recorded. How were the recordings used?

Some individual rankings were reportedly excluded because they were â€œinconsistentâ€ . What does this mean? How many? How was the decision to exclude made? Was this only for village group participants?

The results are presented as if there were only one group involving 85 participants, but there were 8 separate groups, presumably chosen because they might have different perspectives. Did the groups come to different conclusions? Were public perspectives different to health care workers for example? The HIV group clearly gave different priority to HIV relevant interventions. I would have liked to see more comparative analysis between groups.

The results show that the rankings changed as evidence was cumulatively presented. But what was it about the introduction of additional evidence that led to changes â€œ was this apparent from the recordings of discussions? Was any analysis of the recordings undertaken to explore this? Indeed some of the explanations posited in the discussion might have been tested by analysis of the recordings. For example, it is stated that the priority for TB treatment might have been influenced by the high level of present coverage â€œ was this not apparent or testable from the recorded discussions about ranking.

Some of the conclusions donâ€™t derive from the study â€œ e.g. the need for capacity building to summarise the evidence base.

------------------------------------------------------------------------------------------------------------------------------------

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

The title needs revision â€œ it doesnâ€™t capture the key issue of differences between policy makers and public engagement â€œ hence doesnâ€™t tell the potential reader enough about the focus of the paper.

------------------------------------------------------------------------------------------------------------------------------------

Discretionary Revisions (which the author can choose to ignore)

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: No

Declaration of competing interests:

'I declare that I have no competing interests'