Reviewer's report

Title: Economic evaluation of an experience sampling method intervention in depression compared with treatment as usual using data from a randomized controlled trial

Version: 0 Date: 14 Dec 2016

Reviewer: Linda Davies

Reviewer's report:

This paper addresses an important topic in current health care and overall the evaluation appears well designed and implemented. More detail in places would help the reader understand the study and be able to assess its robustness and relevance to their own setting.

Are the methods appropriate and well described?

1. How robust is the sealed envelope allocation procedure?

2. What is the rationale for a 32 week follow up schedule and why were follow up costs and QALYs only collected at weeks 20 and 32. There may well be good reasons for this, but it would help the reader if these were explained in the methods section and/or considered in the discussion section.

3. I assume that the trial start date pre-dated the launch of the EQ-5D-5L. It would be good if the authors confirmed this or gave a rationale for using the EQ-5D-3L if this is not the case.

4. An EQ-5D value set is available for the Netherlands, so it would be useful to provide more detail about why this was not used in this study. Would it make much difference to the conclusions? A sensitivity analysis could assess this question.

5. Service use is valued using a range of unit costs and then standardised to 2012 prices. Looking at the references this suggests that some of the unit cost data are older than 2012
- are there more recent unit costs that could be used? If not, it would help the reader if this was stated as part of a rationale for choice of price year. Are there any changes in the organisation and availability of services that may affect the relative costs of different services that may affect the conclusions? There may not be, but it would be useful information for the reader.

6. It is good to see a thorough description of how the willingness to pay thresholds were set. A bit more detail to support the last sentence in that section would be useful to explain what the 0.41 and 0.7 actually measure, how they were derived and where the participants in this study sit in that range.

7. There seems to be a contradiction in the statistical analysis section. The authors state that 'Coefficients obtained from regression analysis correcting for baseline costs could not be used in the present data, because the data failed to meet the assumption of a normal distribution of the cost residuals,' then later state that 'Because residuals in the analysis were not normally distributed, non-parametric bootstrap resampling techniques were used to explore sample uncertainty around estimates of the cost-utility and cost-effectiveness analysis, using the original data of the three groups.' Statistically there may be good reasons for this, but for readers who are not statisticians it would be good to rephrase or clarify these points. I think you go some way to explaining the issue in the discussion, so either moving some of the text from the discussion to the methods or signposting the reader may be useful.

8. Apologies if I missed this, but what characteristics were included in the regression models and why/how were they identified? I think it would be good to report this in either the methods or results of the main text.

9. I would like to see a list of the sensitivity analyses conducted in the methods section to make this information easily accessible to readers. Details about the how and why are fine in the supplement, but more detail would help readers assess the scope and rationale of the sensitivity analyses and whether they are robust or relevant to their setting.

10. The methods used to impute missing data are well described. It would be good to provide the reader with more information about the rationale for using these methods would be useful, either in the methods section of the discussion. In particular, multiple imputation is widely used in clinical and economic trials. Why were the methods used here thought
to be superior? What were the limitations of the approach used and what are the implications for the conclusions?

Results

11. I know it is a common way to report results, but I wonder whether to would be better to report results where there is no statistically significant difference more neutrally. For example you state that: 'ESM-I participants were more often in symptomatic remission compared with control group participants, although this difference was not statistically significant (OR=2.65, p=0.12);' a more neutral statement would be: 'That there is no evidence whether ESM-I participants were in symptomatic remission than control group participants (OR=2.65, p=0.12);'

12. Table 2 seems to be missing some headings in the version of the paper I have.

13. The cost effectiveness results are reported clearly. It would be useful if you added some interpretation of what the ICER and probability of cost effectiveness mean. For example what is the ICER and unit of unit of health benefit gained?

14. Similarly the results of the cost utility analysis would benefit from some interpretation. What is the ICER? If ESM-I is 44% cost effective at the a priori WTPT, does this mean that it is not cost effective (since the probability is below 50%)? The readers may not be familiar with the methods used and how to interpret them, particularly when 3 alternatives are compared to each other. I note that you do address this to some extent in the conclusion. It may be useful to introduce the issues in the results section.

Does the work include the necessary controls?

Yes. The work includes two controls which appear to be appropriate to the research aims and objectives. Further information about what usual care included would be useful - to help decision makers assess the relevance of this control arm to their setting. It also helps readers assess the validity and relevance of the intervention and control groups.

Are the conclusions drawn adequately supported by the data shown?

I'm not sure the key findings reported in the discussion section are supported in the results reported in the previous section.
1. My reading is that there was no evidence of differences in costs or clinical outcomes. So the only measure where the difference was statistically significant was the QALY.

2. I'm not sure from the methods reported and the results presented whether the CEACs for the cost utility analysis demonstrated that EMS-I was cost effective. It would help if you described the probability of cost effectiveness required.

3. I'm not sure that the results reported in this paper support the statement that: 'However, results presented in both papers pointed in the same direction and we may tentatively conclude that ESM-I is effective in reducing symptoms.' It may well be the case, but it would be good to have more detail about how this conclusion is drawn from what appear to be non-significant results.

**Are the methods appropriate and well described?**
If not, please specify what is required in your comments to the authors.

No

**Does the work include the necessary controls?**
If not, please specify which controls are required in your comments to the authors.

Yes

**Are the conclusions drawn adequately supported by the data shown?**
If not, please explain in your comments to the authors.

No

**Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?**
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

**Quality of written English**
Please indicate the quality of language in the manuscript:

Acceptable
Declaration of competing interests

Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I have was a co-applicant on an unsuccessful grant application (2014) with one of the authors of this study (Jim van Os). I have not worked with Professor Os before or since.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal.