Author’s response to reviews

Title: A randomised trial and economic evaluation of the effect of response mode on response rate, response bias, and item non-response in a survey of doctors

Authors:

    Anthony Scott (a.scott@unimelb.edu.au)
    Sung-Hee Jeon (shjeon@unimelb.edu.au)
    Catherine M Joyce (catherine.joyce@monash.edu)
    John Humphreys (ohn.Humphreys@monash.edu)
    Guyonne Kalb (g.kalb@unimelb.edu.au)
    Julia Witt (witt@cc.umanitoba.ca)
    Anne Leahy (anneel@unimelb.edu.au)

Version: 3 Date: 31 May 2011

Author’s response to reviews: see over
Response to referees comments.

*Our response to referees comments are shown below in italics*

**Referee1.**

Major Compulsory Revisions

1) It is not clear how results are analyzed. Individuals are randomized to one of three arms; online, paper or mixed. However, in all three arms, respondents have the option of completing the survey either online or on paper. So there are really two different levels operating here – contact protocol and mode of completion. It is likely that each level will work differently on the different outcomes. For example, contact regime would impact unit response rate whereas mode of completion would impact item response rate. Please clarify for the reader the distinction between these two levels and how you analytically handle cases that cross modes.

2) In some respects, all three of the contact regimes represent mixed mode designs. In all instances, the individual has the choice of completing the survey either online or in paper. What differs is the ease of each completion mode across contact regimes (i.e., in paper the individual has a paper copy whereas in online they would need to take another step of either requesting the paper survey or printing it out). Please discuss in light of table 5. What you have called mixed mode could be called sequential mixed mode and what you have called paper mode could be considered simultaneous mixed mode. Further, in the online mode you require respondents to retype an address into a computer – this requires more from a potential respondent than would an online mode that contacted individuals via email. This should be discussed.

*We have changed the descriptions of the three arms of the trial (see page 7 and Figure 1), and the title of the paper, to more accurately reflect what we were comparing. We now have two mixed modes as you suggest, simultaneous and sequential. We justify more clearly in the introduction our motivation for choosing these three modes. We felt that, based on previous literature where doctors exhibited preferences for one mode over another, that it was important for doctors to choose the mode that suited them best, within each arm of the trial. We analyse the data by intention to treat, which is standard in the randomised trials literature.*
3) It would be informative to compare responses to the first mailing and the second mailing across modes. For example, the first contact in the online and mixed modes are the same. How do responses differ just to the second contact in these two groups?

We agree this would be interesting. However, we feel that this is beyond the scope of the paper and would require many more tables and associated text that repeat the analysis for the second mailing.

Minor Essential Revisions

1) Abstract, results: Include the overall response rate and/or the response rate of one of the modes to give the differences scale.

Done.

2) Background, paragraph 3: You state that the “literature is large” please specify the literature on what. Also in this paragraph clarify if the cited studies pertain to doctors or the general population.

These issues have been clarified (page 4-5).

3) Methods, paragraph 1: Please clarify how the current study fits into the MABEL study and how the cost results are extrapolations to the full study.

This has been clarified in the methods section (para 1).

4) Methods, paragraph 1: How are the references [3,8,25] supporting the preceding statement?

These have been deleted.

5) Methods, paragraph 1: How complex was the process to logon to the secure website? This could impact the response rate in the mode and should be discussed in the discussion.

This has now been detailed in the methods section (page 9), and is also mentioned in the discussion.

6) Methods, paragraph 1, last sentence: It is not clear which arm this statement pertains to. It is also not clear if a second paper survey was sent in the second mailing within the paper arm.

This has now been clarified in the text and Figure 1 as noted above.

7) Methods, paragraph 2: How long was the survey in total?

This is now mentioned in the methods section (page 9).

8) Methods: You assess nonresponse bias by comparing to the overall
population characteristics of doctors in Australia. Please discuss the quality of these data. Are there any potential sources of error here? Also, do you have these data on an individual level? If so, it could be interesting to directly compare characteristics of responders compared to nonresponders.

The quality of these data are now discussed in the methods section (page 10). We did have this data at an individual level, and this was the basis of comparison with the population. We felt that comparison with the whole population was more useful to assess representativeness. In any case, comparison with non-responders produced very similar results, though these were less precisely estimated.

9) Results, paragraph 2, penultimate sentence: Who is the comparison group?

This has been added in (page 13).

10) Results, paragraph 2, last sentence: The introduction of the incentive here is confusing. Was it offered in all modes? This detail may be better placed in the methods.

This has now been mentioned in the methods section and clarified that it was given in all modes (page 9).

11) Discussion, paragraph 2: It is asserted that doctors “prefer” one mode over another. This is not directly assessed and should be qualified as such.

It is assessed in that doctors were able to choose which mode to complete under each arm of the trial, thus revealing their preference for paper or online completion. We have reworded this slightly to aid understanding by replacing ‘prefer’ with ‘choose’ (page 17).

12) Discussion, paragraph 3: While more younger doctors may drive a switch to online technologies, this is speculative and many alternative mechanisms exist, such as the proliferation of smart phones, for example.

I would not think it desirable to fill out such a survey on an iPhone for example, compared to a large screen.

13) Discussion, paragraph 4: It is not clear how costs will vary across geographies exclusively. Won’t they also vary across institutions? Please expand on this.

This has been clarified (page 19).

Discretionary Revisions

1) Background, paragraph 5, sentence 3: Suggest changing “increase in response rate” to “decrease in nonresponse bias.” As you state earlier these are not always linearly related and as such, nonresponse bias is really a more important indicator of data quality.

We have added ‘or other outcome’ in parentheses after ‘increase in response rate’ (page 6).

2) Background, paragraph 5: Consider citing James, et al. (2011). Health


3) Methods, paragraph 3: It seems unnecessary to me to include which investigator did the randomization, organized the data, etc.

This is necessary in terms of being able to assess potential bias in the trial design and being clear who was blinded to group allocation.

4) Results: Table 1 is essentially a check of the randomization. I think that the table could be omitted and you could just state that the randomization “worked.”

We have left this in, again as part of checking the design of the trial.

5) Results, paragraph 5: You discuss item response rates for the different sections. It may be useful to discuss where in the survey the sections were located. Was there overall attrition or just skipped questions/sections?

The order has been noted in the text (page 14), and is the order that appears in the tables. Overall attrition is non-response, whereas skipping questions is item non-response. Here we examine the latter.

6) Discussion, paragraph 2: The portability of the paper survey is mentioned as a potential benefit of this mode. Increasingly, online modes are also portable. This could be useful to include in the discussion.

This has been noted in the discussion (page 17-18).


Minor issues not for publication
1) Abstract, methods: First sentence is not a complete sentence.
2) Methods, paragraph 3: States is spelled wrong.
3) Methods, paragraph 4: It seems redundant to me to say mean response rates.
4) Results, paragraph 5, sentence 2: I believe you mean “response” not “question”

The above issues have been dealt with.

Referee 2.

I would strongly recommend that the authors provide more definitions upfront to the concepts discussed. For example, on page five survey response bias and item response bias are both discussed without prior introduction or definition.
We now include brief definitions on page 5-6.

Background/Second Paragraph: I would recommend that the authors refer to the literature as a “growing body,” as there remains significant gaps in our understanding of how to best improve physician participation.

This has been clarified – pages 4-5.

Second sentence of the same paragraph (discussion of mode interventions) should include a reference to mixed mode or some combination of paper, internet or interview.

This has been altered in the light of the previous referee's comments (see above).

Last sentence of the same paragraph should reference “perceived” relevance.

Done.

Minor Essential Revisions:
The first paragraph should be expanded to include appropriate definitions as well as implications for lower or declining survey response rates of physicians (compared to general population).

Implications have been added to this paragraph. Definitions have been added as described above.

In the first sentence, data should be referred to as plural.

Done.

Background/Third paragraph: Make explicit that the two studies examining web-based surveys are non-physician studies. There is also a body of literature on physicians’ use of the internet generally that may apply to this discussion.

These parts have now been re-written making this clear (page 5).

Page 7/top paragraph: Last sentence is somewhat confusing, as it is not clear if it applies to each respondent mode or to one mode in particular.

This has been re-written.

Discussion: Would recommend that the authors re-order discussion to follow presentation of results (i.e., response rates first, cost last).

We think that the headline results are those for cost-effectiveness, and so these have been presented first.

Major Compulsory Revisions:
The primary concern with the literature review is that the authors utilize a number of studies that are not physician surveys, without making this clear. Evidence already suggests that physicians differ from the general public in terms of survey
demand, response behavior, and appropriate mechanisms to improve participation. There is also a small body of evidence indicating that there may even be differences among health care providers across each of these domains. Where evidence is not direct, it should be noted with explanation provided as necessary.

This has now been re-written (pages 4-5).

The authors do not cite the full body of literature on impact of survey mode on either response rates or data quality. Dillman has published extensively on this issue. Other authors include, but are not limited to, Beebe, de Leeuw, McGrath, and Johnson. This has implications for the introduction as well as the discussion/conclusion.

We have now added these and other more recent references.

Page 11/top paragraph: This is the first time that a monetary incentive is mentioned. Would recommend that these cases be excluded from the analyses, as it is unclear whether it is mode or the incentive that is driving the likelihood of survey cooperation (which is the subject of that sentence)

We have now mentioned this in the methods section and highlighted that the incentive was offered to doctors in each mode (as detailed above).

Referee 3.

Major Compulsory Revisions

Discussion:
- The Discussion needs a more in-depth approach. Main findings for the four outcome measures should be compared to existent literature and discussed.

This has now been done in more depth.

- The discussion should include more information about study limitations. The main limitation is response rates. The last paragraph states: "but for large national surveys of doctors...the mixed mode seems to be the preferred option". Is this a valid statement when only 20.7% responded? This should be discussed and reflected on. Even though other criteria are acceptable, can surveys with 20% response rate (or 13% for the online group) be trusted? I would also like elaborations related to why the study ended up with more than 80% non-respondents (for instance by reference to the second paragraph of the Background section). Also, might the low response rates affect main findings?

We highlight clearly in the introduction the issues surrounding the often erroneous perception that a low response rate is necessarily indicative of response bias. This may not be the case, and was not the case in MABEL (at least for some observed characteristics). We explicitly examine this issue in table 6, and have added some discussion of this to the Discussion (page 18).
- There were no assessments regarding non-response bias on substantial variables. Was there any difference on main survey variables for the three groups? Could any of the variables be compared to external data sources as a form of validity testing?

*There are many variables in the survey which will be used as primary and secondary outcome measures. Some of these are part of the longitudinal element and so cannot be examined with just one wave (e.g., mobility of doctors). This is beyond the scope of this paper where we are not focusing on the representativeness of primary outcomes, but of the survey more generally.*

- What is the external validity of this study? (other populations, countries). How did this study advantage the knowledge in survey methodology in general, or within the health services research field? What are possible implications of the findings? Any recommendations for research or the practice field?

*This has been added into the discussion.*

- Minor Essential Revisions

**Introduction:**
- The authors write that there are relatively few published studies on the use of the internet for conducting questionnaire surveys. Then they refer to two meta-analyses with a large number of studies included. Other sources also confirm that a range of internet surveys have been published (see for instance http://www.websm.org/, and the special theme in Public Opinion Quarterly in 2008), so I think the statement about few published studies should be rephrased.

*This has been rephrased.*

- One of the referenced reviews show that two study features (population type and follow-up reminders) contribute to the variation of response rate differences between mail and web surveys (Shih & Fan 2008). Medical doctors seem to prefer traditional mail. Could this be used as basis for a hypothesis about response rate in the present study? And based on existing literature, what about hypothesis relating to the other outcomes?

*The introduction has now been re-written to more clearly indicate the motivation for the modes chosen in the trial (page 7) and to indicate specific hypotheses (page 8).*

- I think the authors could motivate the study aim a little bit more, for instance by arguing why multiple outcome measures are important and why these outcomes were chosen (with reference to existing literature and research gaps).

*Again, this has justified more clearly in the introduction, as detailed above.*

**Methods:**
- The description of the three groups shows that all groups actually were mixed, with a different combination of mail and online. Why were these combinations chosen? The following report could be relevant to read and cite regarding the

Done – reference #14 cited on page 5.

- The title seems inaccurate since all groups actually were mixed.

The title has been changed.

- What is AMPCo? Why did they conduct data collection? What did the survey letter say about who was responsible for the survey?

This is clarified in the methods section (page 10).

Results:
- The response rates in the three groups should be more clearly presented.

These are presented in Table 2.