Author's response to reviews

Title: Response rates in postal surveys of healthcare professionals between 1996 and 2005: An observational study

Authors:

Julia V Cook (J.V.F.Cook@ncl.ac.uk)
Heather O Dickinson (heather.dickinson@ncl.ac.uk)
Martin P Eccles (martin.eccles@ncl.ac.uk)

Version: 2 Date: 22 July 2009

Author's response to reviews: see over
Dear Editors

We thank the reviewers for their helpful comments and have addressed each point below and made the necessary revisions to the manuscript for resubmission. We look forward to your response.

Regards

Dr Julia Cook

Response to reviewers comments:

1 Reviewer's report KMM
Discretionary: it would be of interest to have more detail reported in terms of the number of reminders employed by studies and whether or not a point is reached at which subsequent reminders fail to have any additional effect - i.e. what is the optimum number of reminders?
As requested, we have now reported in more detail the effect of reminders (see Table 2). However, we found no significant difference between the effect of 1 reminder, 2 reminders and 3-5 reminders, so these categories were collapsed in multivariable analyses. This has now been clarified (see page 7, line 13 and page 9, lines 18-19).

2 Reviewer's report EVDK
Minor Essential Revisions
1. The objectives are focused and clear, however the operational definition of the primary outcome measure is not clear. Logistic regression was used to assess the effect of various characteristics on response rate. It’s not clear how response rate is categorized into a binary variable for logistic regression. Therefore a clear definition of how the primary outcome variable was measured would be helpful.
This has now been clarified - see page 7, lines 1-4.

2. The population studied is relevant to most healthcare professionals but I
believe it could be made to be more relevant. A nursing-focused database was not searched, in spite of nurses being one of the target groups. Some, but not all nursing papers are published in Medline; however Cinahl is the primary nursing database. This issue should be addressed or at least acknowledged as a limitation.

We have acknowledged that we did not include a primary nurse focused database in the discussion (strengths and weakness section page 12, lines 15-16).

3. Break down study sample size into smaller groups. The category <1000 includes 254 of the 350 studies. This may have been done for statistical purposes; however, as indicated by the large number of studies falling into this category, the majority of surveys in healthcare are much smaller in scale than the large population-based studies. Therefore, more granularity in terms of sample sizes smaller than 1000 would increase the relevance of these findings for the majority of HC professionals and researchers embarking on a survey.

As suggested, we now present descriptive statistics by smaller sub-groupings of number of survey participants (see Table 2). However, adjacent sub-groups in which the response rate did not show any statistically significant difference were amalgamated in the final multivariable model. This has resulted in size groupings that are different from those in the original version of the paper.

4. How were mixed survey designs handled? For example:
   a. Studies that included both a mailed and electronic component.
   b. Studies with telephone follow-up

   A telephone reminder was included as a single reminder. There were variations in how much information regarding reminders was given in the original surveys. Some did not specify what type of reminder was sent (i.e. telephone or postal or electronic).

   We have now clarified this - see Table 1

5. Did more than 1 assessor review the studies to be included?

   This has now been stated - see page 6, line 4.
6. On page 7, 2 large studies (1 US and 1 Canada) were excluded due to undue influence. It would be interesting to discuss how these 2 studies compare to the final results of this analysis. Do they support the results?

We have now performed a sensitivity analysis, including each of these studies in turn in the final multivariable model (see page 10, lines 5-8).

7. Which countries are included in “others”? This represents 138 studies and countries outside of UK/Ireland and USA cannot assess the relevance of these studies to their setting. Could the classification be made more relevant to others, e.g., Europe, North America, UK/Ireland, etc.

As suggested, the breakdown of the other countries is included in the description of studies on page 8, lines 21-24, and in the revised Table 2.

However, in the final multivariable model, we grouped together countries that showed no significant difference in response rates. This has resulted in groupings that are different from those in the original version of the paper.

Discretionary Revisions

8. One third of the studies (108) did not provide information about length, yet this group had the highest response rate. Consider including this “unknown” group in the bivariable and possibly multivariable analysis, rather than excluding it from the analysis.

As suggested, we have now included this “unknown” group in the univariate and multivariable analysis (see revised Table 2 and Table 3 and changes in text on page 7, lines 14-16).

9. Was the quality of the studies assessed in terms of sampling strategies – convenience sample versus random sample?

No, we were limited to assessing features of study design that were ubiquitously reported in the original papers.

3. Reviewer’s report OAB

- Major Compulsory Revisions

1. Background:

a. The authors have chosen an important research topic. I think the parts about estimating response rate and frequency of assessment of potential for
non-response bias are important contributions. However, in my opinion the study does not include sufficient data to answer all the research questions raised. The part about factors associated with higher response rates lacks several of the most important predictors for response rate (i.e. monetary incentives, recorded delivery, importance of survey topic), while the part about changes over time compared to 1986-1995 lack credibility since there are several substantial differences between the current study and the study by Cummings et al. (2001). The current study provides adequate data to answer two of the research questions (estimate of response rate in the period, frequency of non-response analysis), and in my opinion these issues should be the primary research questions in the study.

We have made the following changes to meet all four study objectives.

Objective 2: To assess whether response rates have fallen by comparison to Cummings et al.

We have now compared a simple un-weighted mean response rate in doctors which is directly comparable to the results from the Cummings et al, study - see comparison with previous studies, page 10, lines 11-14.

Objective 3: To explore the influence of multiple factors associated with higher response rates.

We have included our reasoning behind the choice of factors potentially associated with higher response rates (see page 6, lines 11-17). Whilst these have been extensively researched in the general population, their applicability to healthcare professionals is far less certain (see revised background section). The choice of factors associated with higher response rates were limited by a lack of reporting in the studies, for example delivery systems were seldom reported. Furthermore, the assessment of saliency of survey topic in this retrospective design would require at least contact with study authors and is impractical. These factors have been acknowledged in the discussion (see page 12, lines 22-25).

b. The background is short and not very informative. It does not give an adequate review or description of relevant research for each of the four research topics, and it lacks important references. For instance, the highly relevant article by Asch et al. (1997) was not mentioned at all ("Response
rates to mail surveys published in medical journals” (J Clin Epidemiol). The authors should shortly describe relevant research, at least for primary research questions, and connect their own study to this knowledge base.

We have made considerable changes to the background section. This includes reference to the study by Asch et al, 1997 and more information on the current literature relevant to the study objectives (see page 3-4).

2. Methods:
   a. Sample size: It looks like the sample size calculation was constructed to give a basis for testing against the response rate in the study by Cummings et al. Therefore, I’m surprised about all the differences between the current study and the study by Cummings et al. (i.e. search strategy, inclusion- and exclusion criteria, data abstraction, sample size). These differences directly affect comparability. For instance, how can the two studies be comparable when one of the studies only included physicians and the other included physicians, nurses, and allied health professionals? According to Asch et al. (1997) physicians have significantly lower response rate than other groups, and the current study also found (insignificant) differences i.e. between doctors (59%) and nurses (50%) –

   We felt that there were methodological flaws which precluded precise duplication of the methods utilised in Cummings study – these are well explained in the paper (the influence of study size and clustering effects, see page 7, lines 6-11). However, we have now calculated a simple unweighted mean in doctors only allowing direct comparability with Cummings et al (see results page 10, lines 11-14). The study by Asch et al, included many non-health care professionals in the analysis, over 50%, and did not show any difference between individual health professional groups (doctors were significantly lower responders when compared to all other groups) - see revised background section).

3. Discussion:
   a. Are response rates declining: I find this paragraph problematic. As mentioned above there are several important differences between the two studies. For instance, the study by Cummings et al. only included physicians (response rate 61%), while the current study included physicians, nurses, and
allied health professionals. The current study found an average response rate of 57%, but Table 1 shows that the response rate for doctors was 59%, clearly not significantly different from Cummings et al. However, my main point is that the two studies are not comparable because of all the differences mentioned in point.

We have clarified the comparability issues between our study and that of Cummings et al. The response rate of 59% for doctors (in Table 2, 4th column) is the median not the mean and therefore cannot be directly compared with the mean response rate of 61% reported by Cummings. The mean response rate in our study for doctors only is now included (see results page 10, lines 11-14).

2a. I also find it strange to present results for one of the main research questions in the Discussion. We have moved the comparison with Cummings’ results to the Results section.

b. Factors that influence response rates: the current study does not really bring anything new regarding factors to increase response rates. In my opinion, this part of the study is weak since it lacks some of the most important predictors. Current research about reminders concerns how many reminders (not reminders vs. no reminders), and as the authors point out it is difficult to interpret the two other factors identified (larger studies, US studies). I believe this research issue should be of secondary interest, and the weakness related to the lack of relevant predictors should be more explicit. If the authors maintain this issue as a primary research question, the literature should be better reviewed in the Background. It should be clear why the authors chose country, size, publication type etc. and not other well documented initiatives (i.e. monetary incentives, recorded delivery, importance of survey topic), and the consequences of this choice.

We have addressed the issues regarding Objective 3: To explore the influence of multiple factors associated with higher response rates in response to point 1a.

We have i) improved the background section as suggested.
ii) clearly stated the reasons behind the choice of variables studied (see page 6, lines 11-17) and why some well known predictors of response rates from intervention trials were not assessed (see page 12, lines 22-25).

iii) altered the groupings of reminders, country and number of survey participants.

c. Strengths and weakness: I would consider integrating these points into the other paragraphs. As mentioned, I think the main weaknesses relates to the multivariate analysis (lack of several important predictors), and the comparison with the prior 10-year period (several substantial differences between studies).

Standard format for a discussion includes a section on strengths and weakness. We have added to this section based on all the reviewers' comments and feel it should be included in the present format.

4. Conclusions:

a. In my opinion, this study has not shown that response rates appear to be declining, see point 4a. We have now included a more comparable analysis with that of Cummings et al. We have confirmed a small but statistically significant decline in response rates among doctors over the period 1996-2005 but have been cautious to interpret the clinical significance of this finding (see discussion, page 12, lines 1-8).

Minor Essential Revisions

5. Results:

a. I’m unsure about the value of Figure 1. Predictions are based on an inadequate regression model lacking several of the most important predictors for response rate. In addition, the confidence intervals are so wide that they hardly give any meaning, and they all overlap (does this imply no significant differences in predicted response rate between the groups?). As suggested, we have removed the Figure.

b. Table 2 includes a lot of information from different types of analysis. I wonder if it would be easier for the reader if this table was split into at least two tables; i) descriptives; ii) univariate/multivariate analysis (maybe even a
third table to allow for more information about the multilevel model). P-values should be included.

As suggested, we have divided Table 2 into two tables, one presenting descriptive statistics and univariate results (Table 2) and another presenting multivariable results (Table 3). We have now included p-values in both tables.

c. Table 2: what was the ICC in the multilevel model?
This is now reported in the text (see page 10, line 4).

d. Page 8. What do you mean by “After allowing for these associations, most of the remaining unexplained variation in response rates (84%) was between studies and could not be ascribed to sampling variation within studies”? A clearer and more “educational” description of the results from the multilevel analysis would be helpful.
This is an interpretation of the ICC. We have now clarified this section.

e. I believe it would be very useful to include a figure showing the response rate for each year in the period 1996-2005. Regardless of statistical uncertainty, this would give valuable information for the question of changes over time in the period.

It was not the objective of our paper to study the trend over years and the study was not statistically powered to assess a time trend; our objective was to assess whether the response rate was lower in the ten years from 1996 to 2005 than in the previous 10 years (1986 to 1995). We therefore feel that addition of such a figure would distract from the main focus of our study and, due to its statistical uncertainty, have the potential to mislead the reader.

Discretionary Revisions
6. Conclusions:

a. I am not sure what is meant by the final statement to journal editors: “…and make no attempt to understand the implications of this”.

We have clarified this statement see page 15, final line.