Author’s response to reviews

Title: Accuracy of responses from postal surveys about continuing medical education and information behavior: experiences from a survey among German diabetologists

Authors:

Sven Trelle (trelle@uni-hamburg.de)

Version: 2 Date: 7 Jul 2002

PDF covering letter
**Author’s response to reviews**

**Title:** Accuracy of responses from postal surveys about continuing medical education and information behavior: experiences from a survey among German diabetologists

**Author:** Sven Trelle (trelle@uni-hamburg.de)

Dear Dr. Sawyerr and Dr. Kaner

Enclosed please find the revised manuscript. I would like to thank Dr. Kaner for her comments.

**Abstract**

1. (portion of inaccurate responses): "notable" instead of "substantial" (page 2, line 17)

**Background**

1. (change which was not requested): (p3, l18): "Quality Control"

**Methods**

1. (questionnaire - provenance): I rewrote the beginning of this section (p5, l3-9):
   "The data used for this sub-analysis was collected by an explorative survey about information management and CME habits (for details see[22]). For this survey a new questionnaire had to be developed. Initially a preliminary questionnaire was developed considering three already published surveys [19, 20, 23]. It was discussed with members of the research group and sent to experts requesting comments. After incorporating these comments the questionnaire consisted of 92 items."

2. (sample size calculation): I rewrote and rearranged the description of the sample size calculation (p5, l17-24): "The sample represented 29% of all 1585 diabetologists in the database. Sample size was calculated with regard to confidence intervals for estimated population frequencies (95% CI): a maximal margin of error for proportions of ± 6.25% for questions answerable dichotomously was considered narrow enough (i.e. the maximum width of the 95% CI for proportions should be 12.5% for questions with only two response categories e.g. yes/no). Given the population of 1585 diabetologists this required a sample size of 213. Response rates of prior surveys ranged from 50% to 70% which results in a sample size of at least 416 persons. For technical reasons it was not possible to draw a random sample. Therefore the sample was determined by the first figure of the zip code (code 1-3)."

3. (questionnaire - test criteria): Since this survey was an explorative study the possibility of assessing test criteria is limited. I am perfectly willing to write a section about this if you think it is necessary or helpful. The other article about this survey is published in the meantime [22]. Strictly speaking it is not possible to make any statement about the objectivity of the survey since I did not vary the survey method (e.g. variation of questionnaire distribution). Although not formally assessed I presume that the objectivity of the analysis of the questionnaire is relatively high. Interpretation of the data is somewhat subjective since there are no normal values for comparison. Given these uncertainties it is actually difficult to make any statement about the other test criteria. Nevertheless the analysis in this article should provide an indication of the reliability in terms of internal consistency (Retest-reliability or paralleltest-reliability was not tested). Face or content validity seems relatively high (see
questionnaire development). Construct validity is very low for one question as published recently (BMJ 2002; 324: 950-951). For the other questions no data are available (including the other previously published questions).

4. (validity of the method): Since there are no specific comments on the other variables I restrict my response to the time variable and the fictitious term. I had no idea how to correct the time-variable-part therefor I deleted it (p6; p8, p18). It may be possible that respondents confused the fictitious term with McNemar’s test or that they thought the researchers were confused although I do not think this is very likely. The term McNemar-Quality-Scale sounds quite different from the term McNemar test (German: McNemar-Qualitäts-Skala vs. McNemar Test). Moreover somebody who knows a statistical test would (probably) know the term Alpha-Error/Type-I-Error. Furthermore nobody during the development of the questionnaire remarked that their could be a misunderstanding. Nevertheless I added these comments in the discussion (see below).

Results

1. (response rate): This is certainly one of the problems with this survey but I think this point is discussed in depth on page 11 (Discussion - Methodological issues). Nevertheless I added (p10, l16-18): "Nevertheless caution should be applied when generalizing the results of this survey and rates or numbers should be interpreted as a trend rather than at face value."

Discussion

1. (McNemar-Quality-Scale): I added (p11, l22 - p12, l2): "One might argue that positive respondents confused the fictitious term with McNemar’s statistical test or that they thought the researchers had been confused. But this seems not very likely since nobody during the development of the questionnaire referred to this potential problem. Moreover somebody who knows a statistical test would know the term Alpha-Error/Type-I-Error which was not the case in the majority of the positive respondents."

2. (Change which was not requested): (p12, l3-6) "The proportion of inaccurate responses to this knowledge-question should be viewed as a very conservative estimate. A recently published study found that virtually nobody who stated that he/she allegedly understands the technical terms of the questionnaire developed by McColl et al. actually did so [27]."

Conclusions

1. (conclusions venture to far): As I state in the background section surveys are susceptible to various biases and according to the literature some are more important than others. I tested and discussed why the socially-desired response bias in this survey probably did not play a prominent role. Acquiescence or the tendency for only yes- or no-responses are also not very likely biases given the questions (see Appendix). In the discussion it is explicitly mentioned that no definite conclusions about contributing biases in this survey can be drawn (p10, l23 - p11, l2). Therefore I offer three different options in a cautious wording (carelessness, failure in self-perception and - in the revised manuscript - misunderstanding about specific terms).

Changes made in discussion: (p12, l24-25 and p13, l2-4) "Acquiescence or the tendency for only yes- or no-responses as two other potential and important response biases are also unlikely due to the same reasons. Thus it is believed that the inaccuracies in this survey are rather a problem of careless reading/answering or a failure in self-perception/overestimation of competency. Furthermore misunderstanding of questions or about specific terms might also have contributed to the inaccuracies as was shown in a recent study [27]."
Changes made in conclusions (p14, l5-7): "The analysis indicates that it rather seems to be a problem of careless reading/answering of questions, a failure in self-perception or a misunderstanding about specific terms or questions."

2. (portion of inaccurate responses): "was around ten percent" in stead of "lies between five and ten percent" (p13, l3).

References


Acknowledgement

1. I acknowledge Dr. Kaners comments. If she has any objections please change her name to "the reviewer".

Authors’ contribution

1. Sven Trelle conceived, designed, conducted, analyzed, and wrote the study.

Minor linguistic changes were made without marking (e.g. "After incorporating these ..." instead of "After incorporation of these ...").

Yours sincerely,

Sven Trelle