Author's response to reviews

Title: Analysis of clinical uncertainties by health professionals and patients: an example from mental health

Authors:

Keith Lloyd (k.r.lloyd@swansea.ac.uk)
Matteo Cella (m.cella@swansea.ac.uk)
Michael Tanenblatt (mtan@us.ibm.com)
Anni Coden (anni@us.ibm.com)

Version: 2 Date: 23 February 2009

Author's response to reviews:

We thank the referees for their positive and constructive comments. We have addressed these point by point giving details of where in the manuscript we have made changes.

We have submitted 2 additional files:
Lloyd revised.pdf.
Lloyd revised_track_changes.pdf (132K)

The former is our revised manuscript incorportaing all changes. The latter is a track changes version of our manuscript so that you can see where we have made changes.

All page references in this covering letter refer to the track changes version of our revised document.

9736887462475763_comment Anthony Jorm

1. In the Methods sub-section of the Abstract, some details should be given that are currently in the Results sub-section: the number of participants; reference to the PICO format. Also in the Methods sub-section, it needs to be made clear what the linguistic taxonomy refers to (“a linguistic taxonomy of....”).

The abstract has been revised in line with this request (page 2 para2.).

2. In the Methods section of the main text, the number of patients and clinicians needs to be given. It also needs to be clarified how the convenience sample referred to there relates to the DUET database mentioned in the Results. If they are the same, this needs to be made clear.

We have added the number of patients and clinicians to the methods section (page 4 para 3). We have clarified how the sample was derived (page 4 para 3).
3. Given that the P of PICO stands for Patient and all the questions were selected to be about schizophrenia, it is not clear why this element was not always satisfied.

We took the view that merely stating “schizophrenia” was an inadequate description of the patient component of the PICO. An adequate specification required mention of age group or gender.

4. In the Methods it is stated that “discrepancies identified and discussed to agree on definitions. We then assessed inter-rater agreement”. It is unclear whether inter-rater agreement was done on independently made ratings or after there had been discussion between raters.

We have clarified that inter rater reliability was assessed on independently made ratings (page 5 para 1).

5. In the Methods it is stated that “Questions were clustered according to an emergent taxonomy”. This needs to be explained in sufficient detail to allow replication by other researchers. Who did it and what methods were used?

We have expanded the explanation of our methodology (page 5 para 1)

6. On p. 5, it is stated that kappas showed “moderate agreement”. However, on the next page some of the kappas are described as “low”. There have been published standards for giving verbal descriptions of kappas. It might help to consult one of these and use the terms for strength of agreement accordingly.

The strength of agreement is now clearly stated using Landis and Koch’s original classification (page 6 para 1) and (page 6 para 2)

7. Page 6 mentions that the chi-square test could not be carried out because of the low numbers and assumptions not being met. Exact tests are widely available these days (e.g. in SPSS) which overcomes this problem.

The main reason for omitting the patient questions from this part of the analysis forms our answer to the reviewer’s point 6. Therefore we have deleted the sentence referring to invalid chi square assumptions (page 6 para 2)

8. On p. 6, it is unclear what the chi-square tests are comparing. This should be described in the Methods. In particular, the first chi-square test involves all the PICO elements. However, if there was a count of number of PICO elements mentioned, this test would not be the appropriate approach.

We have made an error in the result reporting. The first chi square value refers to the P and not to all the PICO elements. We have amended the sentence as
Naturally occurring questions that involved “what” type queries were the ones most likely to contain identifiable PICO elements. Strong associations were observed between the What category and the “Patient” component of the PICO (x² = 33.31 p<0.0001) I (x² = 27.52, p<0.0001) and O (x² = 23.97 p = 0.001). (page 6 para 3 and page 7 para 1)

9. In Table 2, it is unclear what denominator the %s relates to. This could be indicated with a total n which corresponds to 100%.

We apologise, this lack of clarity arose from condensing a lot of data into a simple tabular format. We have added a key to the foot of table as follows:

“Key: n = number of questions in that category; % = number of questions in that category / total number of questions x100.” (page 13)

10. Similarly, in Table 4, it is not clear what denominator these %s are of.

We have clarified table 4 so that it is clear that we are referring to column percentages. (page 15)

Minor Essential Revisions

1. On p. 6, use the Greek letter chi rather than x. – done (page 7 para 1)

2. On p. 6, “How” has a capital, but not “What” etc. – all now lower case (page 6 para 2)

3. Page 7, line 2, has missing apostrophe.
   Two apostrophes have been inserted: “However, the team was often able to identify PICO elements present in patients’ and clinicians’ questions.”

4. Page 7, last sentence of paragraph 1, “information…are” should be “information…is”.
   Amended

5. Page 7, last paragraphs, “It is basic” should be “It is a basic”.
   Amended

6. Table 4. I think “30.65” is meant to be “30.6%”.

Amended
Amended

Discretionary Revisions

I have broader conceptual issue with this research that the authors may wish to consider..... I would have preferred to see this type of research carried out in a more realistic way using an actual database like PubMed or Google, with people trying to answer their questions as they naturally would. It may be that people actually do use the PICO elements, but not in one query.

We entirely agree with these observations and are already in the process of carrying out just such analyses for our next publication on this area.

1931721416248525_comment Stefan Priebe

1. The paper appears strangely vague about the precise approach and methods used in the study. The details of how interviewees were approached, how many responded and so on may or may not be relevant for the interpretation of the findings, but they should be reported to some extent so that the reader has a clearer picture of how the study was conducted. The authors say that they used a convenience sample. Still, who was approached, how and where, and how many responded? Was there a particular strategy involved?

We have expanded the methods section so as to specify our recruitment sources (pages 4 and 5). We cannot specify a response rate as we do not know how many people looked at our web survey but did not complete it. We did address these shortcomings in the final paragraph of the discussion. In view of the reviewer’s concerns we have expanded this section also (page 7 para 3 and page 8 para 1).

2. What was the context for the whole study? For example, the authors mention that this is a secondary analysis, but it remained unclear to me what the primary analysis might have been. I am sure there is nothing to hide, so I would encourage the authors to be less opaque and provide more precise information.

The data were originally collected to identity treatment uncertainties in mental health and to contribute these to the UK Database of Uncertainties about the Effects of Treatments (DUETs) Specialist Library held by the NHS national library for health. This has been made explicit in the discussion.(page 7 para 3)

3. The whole point of the study is to demonstrate that the complexity of the real world cannot be reduced to the technicalities accommodated by EBM. In the process, the authors report on a linguistic analysis. However, they begin the whole study with a technical term, i.e. uncertainties, that is linguistically imprecise in itself. To me, it looks somewhat sloppy to equate “clinical uncertainties” with
“uncertainties about treatment effects”. The difference between these two concepts alone might underline significant limitations of EBM. Personally, I would not expect elaborate definitions, but (b) some explanation of the significance of the terms and the concepts behind them, and (b) more linguistic precision in using the terms throughout the paper. Why did they use the term uncertainty? I am sure there are good reasons and they should be provided. In the process, the authors report on a linguistic analysis. However, they begin the whole study with a technical term, i.e. uncertainties, that is linguistically imprecise in itself. To me, it looks somewhat sloppy to equate “clinical uncertainties” with “uncertainties about treatment effects”. The difference between these two concepts alone might underline significant limitations of EBM. Personally, I would not expect elaborate definitions, but (b) some explanation of the significance of the terms and the concepts behind them, and (b) more linguistic precision in using the terms throughout the paper. Why did they use the term uncertainty? I am sure there are good reasons and they should be provided.

We have added a definition of uncertainty to the background (page 3 para 2). The term “uncertainty” has a precise meaning within the field of evidence based medicine. We believe this simple definition and accompanying two pertinent references will address the reviewer’s concerns here. We are grateful to him for drawing our attention to this oversight.

Discretionary Revisions

1. I wonder whether the presentation of the results can be reorganised. To me, Table 3 is the starting point for a content analysis which may then lead to a categorisation according to the PICO scheme. As a reader, I would first like to see the material in all its rich diversity before the information is reduced to a small number of pre-defined categories. I would also like to see more description of the material as illustrated in Table 3.

We have limited the amount of information displayed here to keep to a single message in our paper. The idea of a content analysis was indeed discussed by the research team and has been undertaken for the whole data set.

2. I do not see why the focus on schizophrenia is a limitation of the study. I would rather regard it as a strength that the questions relate to a defined illness category.

We thank the referee for this observation.