Author's response to reviews

Title: Communicating projected survival with treatments for chronic kidney disease: patient comprehension and perspectives on visual aids

Authors:

Frances Dowen (francesdowen@doctors.org.uk)

Karishma Sidhu (ksidhu@adhb.govt.nz)

Elizabeth Broadbent (e.broadbent@auckland.ac.nz)

Helen Pilmore (hpilmore@adhb.govt.nz)

Version: 1 Date: 14 Aug 2017

Author’s response to reviews:

Many thanks for the response to our submission regarding our article Communicating Projected Survival with treatments for chronic kidney disease: patient comprehension and perspectives on visual aids, and for the reviewers helpful comments and queries.

Please find below our point by point responses and our amendments in the text. All comments have been addressed and the article amended. We hope this is to the reviewers’ satisfaction.

1. Please add a conclusions section after the discussion which should state clearly the main conclusions and provide an explanation of the importance and relevance of the study reported.

I have done this

2. Please move the declarations section to after the conclusions and ensure that you have the following sub-sections with the appropriate titles:

- Ethics approval and consent to participate
- Consent to publish
- Availability of data and materials
- Competing interests
- Funding
- Authors' Contributions
- Acknowledgements

I have moved this section and included all the subsections. Please let me know if further information is required.

3. In the authors' contributions section please refer to each author by their initials and please also add a statement to confirm that all the authors read and approved the final version of the manuscript.

I have done this.

4. In the ethics approval and consent to participate section please add a statement detailing the consent sought from the participants.

I have done this.

Reviewer reports:

The different types of graphs convey different amounts of information. The Kaplan Meier curve and histogram convey information about survival rates from between 1 and 10 years whereas the pie chart and pictograph only convey information about the 5 year survival rate.

Some of the graphs show 10 year follow up data and some only up to 5 year data. The 10 year follow up data is shown in graph types that more easily carry this follow up ie. Bar chart and Kaplan meier. The questions put to participants are looking at 5 year data, which can be read from all the graphs. In no scenario was a participant asked to directly compare graphs using different survival time data, as they were only shown one type of graph for each of the scenarios listed in the methods. Figure 1 is simply an example of one of the sets of graphs used in one scenario ie. Each participant would only see one of the graph types, not all four.

The pictograph is not a good example of the pictograph because you have one stick man to convey 10 people - and therefore have an issue with showing half a stick man to communicate the difference between the two options.

We agree with this limitation and it is discussed in Discussion on page 15 in terms of not negatively framing data, see additional comment marked in text.

The correct answer provided in the multiple choice question to test comprehension was ambiguous. Options a) and b) state "That people who have a transplant before needing dialysis (after five years on dialysis) live longer". But graph only provides information about the numbers of people who live or die 5 years after dialysis and 5 in 100 more people will be alive 5 years after transplant if they have had the transplant after 5 years on dialysis. So 5 in 100 people will live longer but the answers provided imply that everyone will live longer - which is not strictly correct. The correct answer should therefore be stated as "That more people who have the transplant before needing dialysis are alive 5 years after the transplant".
Option A states that people who have a transplant before needing dialysis live longer, option B states that people who have a transplant after 5 years on dialysis live longer. I agree that the pie charts appeared ambiguous as I had failed to put in the text that both of the pie charts shown in Figure 1 were 5 year post transplantation data and this was stated beneath the pie charts in the survey given to the patients analysed – I have now inserted this comment into the methods. If a participant was shown the pie charts they received both pie charts to interpret. We also agree that the correct answers could be stated as ‘more people’ will be alive after transplant if they have not required dialysis, rather than the approach we used to make a point to patients in informing their choice about live or deceased donor transplantation. See additional paragraph regarding this limitation in the Discussion on p15.

The procedure states that participants were offered an explanation of each graph if requested but they do not present any data about how many participants requested an explanation for each graph type (which would have been extremely valuable data).

All surveys were explained by the researcher. Most patients required an explanation of the graphs as would be done in clinical practice if these tools were used in a clinic setting. The number of explanations required was not recorded but we appreciate that this would have been valuable information in helping to create a standardised accompanying explanation in future. See additional comment p16.

The chi2 test is not appropriate because the observations between the different conditions are not independent of each other - the study was a repeated measures design where each patient received one of each type of graph (independence of observations is a requirement of the chi2 test and each participant must only contribute data to one and only one cell in the chi2).

We have repeated the statistical test and undertaken a Cochran’s Q test which is designed for repeated measures. This is detailed in the Data Analysis section p10 and showed a significant difference in correct interpretation of graphs.

The study is underpowered to show a meaningful effect size. The authors used a sample size which would be sufficient to detect a 5% difference with 90% confidence but they apply a p-value of p<.05 to evaluate significance. As a consequence a potentially meaningful difference in the number of people correctly answering the multi-choice question between the pictograph and Kaplan Meier curve graphs is dismissed as not significant on the basis of a p-value of 0.06.

Following Cochran’s Q test, rather than Chi squared test (on the recommendation of the previous comment), the p value here was 0.0439 therefore statistical significance shown and pictograph more likely to be correctly interpreted than Kaplan Meier.

Peggy Sekula (Reviewer 2): Communicating projected survival to patients is always difficult - but just not limited to understanding. Therefore, to optimize way of presentation is surely of great interest.

The study is nicely conducted. Still, I would like to get some further informations:
I am aware on how GFR is estimated in Europe or in US but how is it calculated in your population (NZ).

The MDRD was used as has been found to be the most accurate in the population. This information has been added to the Methods section.

Similarly, how is ESKD defined? I am used to ESRD (R=renal, KDIGO) instead of ESKD. Obviously, there is no common denomination.

have changed the terminology to ESRD

Table 1: You probably mean “Cause of CKD” and not “Cause of ESKD”

This has been changed (p20)

How were data on comprehension collected? Did the participant complete a questionnaire or did the “member of the research team” record (standardized?) participants' answers?

Was it always possible to distinguish correct from wrong answers? I could imagine that this is not always clear. How did you deal with ambiguous answers?

Participants completed questionnaires in the presence of a member of the research team. If an ambiguous answer was given after explanation of the graphs, it was marked as incorrect. We have added an additional comment p9.

As described, graph types were randomly assigned to scenarios per patient, so each patient received all four graph types. What about distributions per graph types (equally distributed per scenario, same number assessed overall)?

Each participant had four scenarios and four graphs in total, one of each graph type. The same number of each graph type was assessed in the scenarios as there was equal distribution per scenario. See additional comment p8.

In Europe at least, we avoid red and green in graphs because of red-green color blindness. Please comment.

Thank you for this point – additional comments have been made in the discussion (page 15)

Is there any standard reasoning behind 5% difference in sample size calculations? Please specify details for sample size calculation (test, software).

This is a standard difference although we agree that some studies use 10% as clinically significant. The sample size software used was G-Power (http://www.gpower.hhu.de/en.html). I have updated the paper with this information.
Please be more explicit as statements like “Age >65 was associated with a lower frequency of correctly interpreting the histogram (p=0.008).” are somewhat unclear. Please specify such phrases and also in table 2 to what you compared to?

Clarified in text p11. Table 2 compares the ability to correctly interpret a graph type with a certain demographic ie. Histogram interpretation compared in different demographics age found to be significant negative impactor. Title modified to reflect this.

Table 2: please add absolute frequencies and explanation(s) on what hypotheses were tested in footnote.

Absolute frequencies added and footnote added. Hypothesis tested was the likelihood that different demographic variables would be associated with ability to correctly interpret graphs.

Table 3: Please refine presentation to make clear that this is a presentation in the overall cohort of 177 patients by, for example, adding another column for unemployed and another row with totals.

Total column added to table

Table 4: Please add explanation(s) on what hypotheses were tested in footnote.

Footnote added. Hypothesis tested was that different demographic variables would be associated with the likelihood of correctly interpreting all four graphs.

An angle that would nicely add to this study is to also assess understanding and ability to explain graphs by treating physicians themselves. Did you consider that? In my experience there is often a lack. It may worth discussing that.

Yes, this is an interesting point, I have added a comment regarding this on p16.