Author's response to reviews

Title: Risk Communication Formats for Low Probability Events: A study of patient preferences

Authors:

James G Dolan (jdolan787@gmail.com)
Steven Iadarola (stephen_iadarola@urmc.rochester.edu)

Version: 2 Date: 7 January 2008

Author's response to reviews: see over
January 7, 2008

To the editor,

Attached is a revised version of our paper “Risk communication formats for low probability events: a study of patient preferences”, MS: 1301310431153297. In making our revisions, we have tried to address the issues identified by the reviewers about the original manuscript. The changes we have made in response to each reviewer are summarized below. We are grateful to all four reviewers for their helpful and insightful comments.

Reviewer 1, Stacey Sheridan, Major compulsory revisions

1) The objectives are unclear.
   We have clarified the objectives of the study in the Abstract on page 2 and in the Background section at the bottom of page 5.

2) The study’s small convenience sample is a problem unless the study is framed more clearly as a feasibility study.
   The revised study objective clearly identifies this as a feasibility study. As much additional information as we have available about the study population has been added to the description on page 6.

3) Important detail is lacking about measurement.
   A full description of the AHP-based measurement procedure is included on pages 7 and 8. Additional discussion about the use of the AHP to measure preferences and similar concepts has been added to the discussion on page 10. In addition, we have successfully uploaded the computer program used for the study. This will allow the reviewers (and hopefully eventually readers) to examine and actually use the measurement tool we used in this study.
Reviewer 1, Stacey Sheridan, Minor essential revisions
1) The study overstates its findings by implying in the discussion that preferences are equated with effectiveness.
   The distinction between preferences and effectiveness has been clarified.

2) More details about the study scenario would be helpful.
   We have tried to describe the study scenario clearly. Successful submission of the study intervention file should also make this more clear.

3) It would be helpful to know whether the acceptable cutoff for the consistency ratio for AHP was determined a priori or post hoc.
   The consistency cut-off was defined a priori. This has been clarified in the text on page 8.

Reviewer 1, Stacey Sheridan, Discretional revisions
1) It would be useful to have information about comparisons between vertical bars and icon displays with and without the magnifier scale.
   Unfortunately, to limit the duration of the study interview, we did not include un-augmented versions of the bar charts and icon displays so have no data to report.

2) The additional files that accompany this article are unaccessible at the time of this review and should be reposted.
   We have successfully uploaded the originally missing file. We are grateful to the editorial office for their assistance.

Reviewer 3, Jessica Ancker, Major compulsory revisions
1) It is a serious limitation of the study that the authors did not screen the study sample for numeracy or health literacy.

   We did not obtain information about the numeracy and health literacy skills of our sample as this study was intended to be the first in a series of related investigations. Because this was a pilot study, we only collected a
limited amount of data about the study subjects. What we have has been included in the sample description on page 6.

2) The authors do acknowledge that preferred visual formats are not necessarily the most "effective" for risk communication, but this point deserves some additional discussion, as well as a clarification of the meaning of "effectiveness".

We have tried to make the difference between visual display appeal and effectiveness more clear. The clarification of the major objective of the study - to determine whether these new graphs are worth investigating further - should also help explain why we measured preferences in this initial study. We have also added a comment about this issue to the discussion on page 10.

3) In a related topic, on page 10, the authors claim that the study supported the hypothesis that frequency diagrams could enhance communication about low-frequency risks.

This part of the discussion has been deleted.

4) This is an unusual application of the analytic hierarchy process…

The method and rationale for using the AHP to measure preferences has been revised in the methods section on pages 7&8 and in the discussion on page 10.

5) Another point that requires explanation is the use of bar charts and pie charts to illustrate the preference ratings.

The use of linked bar and pie charts to assist with graphical pairwise comparisons is common in commercial AHP software and has seemed to help in previous studies we have done using the AHP. For this reason we kept the same format. In this revision we have not added a specific comment about this for fear it would confuse the main message of the paper. We can include additional comments if the reviewers think it necessary.

6) On a related point, in Figure 5, what is the meaning of the menu window in the upper left-hand corner that says "baseball vs ..." etc.?

The menu window is the navigation menu for the program used to assess the preferences. The baseball reference is to the learning exercise we used to teach the subjects how to do the comparisons. We have added an explanatory comment to the figure legend and included mention of the fact that the subjects did a training exercise judging the relative weights of different balls before they started the main set of comparisons in the
description of the study intervention. We can include additional comments if the reviewers think it necessary.

7) Also in the figures, a limitation to the generalizability of this study is the use of red and green dots in the icon display, given the prevalence of red-green color-blindness among males.

We had no problems with color blindness in this study. We have added the need to develop and refine the characteristics of the newly developed graphs as a limitation and implication of the current study in the text.

Reviewer 3, Jessica Ancker, Minor essential revisions

1) The authors should clarify that with the Bonferroni, the overall alpha was kept at 0.05, not the pairwise comparison alphas. Also, to be a stickler for accuracy, it was the alpha level that was kept at 0.05, not the p value.

The description of the statistical test interpretation has been clarified according to the reviewer’s suggestions.

2) Last line of page 9: The authors should clarify that this increase was an absolute increase, and should also provide the baseline risk to persuade us of their point, which is that this was a very small increase.

The risks of heart attacks associated with rofecoxib and without rofecoxib have been clarified to make the baseline and increased risk more clear.

3) Authors must state what the values in the boxplots are (e.g, median, interquartilerange, range, etc.), and how the outliers were defined (e.g., 3 SD above the mean, etc.). In general, more descriptive legends would be helpful.

Descriptive legends have been added to the boxplot, figure 6, as well as figures 1,2,4, and 5.

Reviewer 3, Jessica Ancker, Discretionary revisions

1) The first paragraph of the discussion is the same as the material in the introduction and is thus redundant.

The discussion section has been almost completely re-written to focus on interpreting the results and not repeating what was said in the introduction.

2) On page 4, last sentence of second paragraph is garbled - perhaps should read "change in *risk* of disease."
The missing word was “risk” – this typo has been corrected.

Reviewer 4, Joan Austoker, Major compulsory revisions

1) *However, it is important for the authors to state where the presented information concerning the reduction of risk in colorectal cancer was derived (are these percentages based on evidence or are these hypothetical??).*

We have added a comment regarding the similarity between the absolute risks and risk reductions used in the study and outcomes associated with currently recommended colorectal cancer screening tests to the description of the study intervention.

2) *It is strange to see the study hypothesis in the statistical analysis section. This should be in the introduction.*

We have clarified the study objective to be an exploratory study to see if further study of these new risk presentation formats is warranted. As such there was no firm study hypothesis. This point has clarified in the methods section.

3) *There is some concern that the study may not been able to detect statistically significant differences. The authors need to include some indication whether they had enough subjects to achieve significance.*

Because statistically significant differences were found among the formats, we did not add a discussion of whether the study sample was large enough to show significant differences.

4) *Results: More information is required in this section.*

We have not changed the result section because it seems to be appropriate in the context of the other changes we have made.

5) *There is no discussion as to how these visual representations would possibly affect decision-making. This needs to be addressed by the authors.*

The discussion section has been almost completely re-written.

Reviewer 2, Brian Zikmund-Fisher, Major compulsory revisions

1) *The most significant concern is that the study design aims to compare different graphs, yet the graphs that are used differ in many aspects other than the format.*
We have added the need to further examine formatting details of the new graphs to the study limitations on page 10. The overall objective the study has been clarified to indicate that the purpose was to see if additional work with these formats, including the very important details noted in this comment, is warranted.

2) The conclusion "that bar charts and icon displays that include magnified views of differences between decision alternatives are acceptable to patients" appears only marginally supportable from this data.

There is a comparison between the new formats and the flow diagram, which is a well-studied and recommended risk presentation format. Unfortunately, to limit the duration of the study interview, we did not include un-augmented versions of the bar charts and icon displays so have no data to report. The issues raised are good ones, just beyond the scope of this initial feasibility study. The restatements of the study objectives and results should make this more clear.

3) The conclusion "that natural frequency flow diagrams may be as effective as other currently recommended risk presentation formats" is also unsupportable with this data.

We agree, this conclusion is not supported by the study data and we have deleted it from the revised manuscript.

4) "Small risks are difficult to display in bar charts and icon displays." Not really, it's just that small risks mean that the colored area is relatively small when compared to the total…

We have tried to clarify the importance of properly conveying and understanding small risks in medical decision making with the rofecoxib example. We believe proper understanding of small risks, as well as small benefits, less than 5% is very important for good medical decision-making and is an area that has been relatively neglected. The reviewer answers his own question about the need to investigate the use of magnified views in his comment when he says: “The traditional uses for ‘magnifiers’ in risk communication have been for situations where the risk is smaller than 1%, a level of detail that neither bar nor icon displays are designed to handle effectively.” This is the level of risk that this work is focusing on (the absolute risk difference in risk of cancer in the study scenario was 1.1%; the absolute risk difference in cancers prevented is 0.9%) and explains the need to investigate new communication formats.
5) *The first page and a half of the discussion does not discuss the results at all but simply summarizes existing research (and would be hence more relevant to the introduction).*

The discussion section has been re-written to focus on the results and their interpretation.

Reviewer 2, Brian Zikmund-Fisher, Minor essential revisions

1) *there's no need to report a p value smaller than 0.001*

We have not changed the p values because this change could cause readers to underestimate the level of certainty we can attribute to the conclusion that the differences we saw in the preferences among the six formats were not due to chance alone.

Reviewer 2, Brian Zikmund-Fisher, Discretionary revisions

1) *Why did the authors include both bar and pie charts to visually represent the preference dimension?*

The use of linked bar and pie charts to assist with graphical pairwise comparisons is common in commercial AHP software and has seemed to help in previous studies we have done using the AHP. For this reason we kept the same format. In this revision we have not added a comment about this for fear it would confuse the main message of the paper. However, we have added an explanatory comment to the figure legend.

2) *I have other concerns about the fact that preferences were elicited as a trinary response and then strength of preference was elicited separately.*

The AHP method used to elicit the preferences has been widely used and has been repeatedly shown to yield valid preferences when validated against external standards. Full discussion of this is beyond the scope of this paper but this issue is addressed in the Gass article, reference 30.