Author's response to reviews

Title: Physicians intentions and use of three patient decision aids

Authors:

Ian D Graham (igraham@ohri.ca)
Jo Logan (jlogan@uottawa.ca)
Carol L Bennett (cbennett@ohri.ca)
Justin Presseau (justin.presseau@alumni.uottawa.ca)
Annette M O'Connor (aoconnor@ohri.ca)
Susan L Mitchell (smitchell@hrca.harvard.edu)
Jacqueline M Tetroe (jtetroe@ohri.ca)
Ann Cranney (ancranney@ohri.ca)
Paul Hebert (phebert@ottawahospital.on.ca)
Shawn D Aaron (saaron@ohri.ca)

Version: 2 Date: 24 May 2007

Author's response to reviews: see over
Dr. Kowalczuk

Please find attached our revised manuscript, “Physicians’ intention and use of three patient decision aids” [MS 2105571003127821], which has been modified in accordance with the reviewers’ comments. In the attached pages, we have addressed each of the points raised by the reviewers.

The manuscript has not been previously published nor is it currently under review by any other journal. It represents original work and all authors listed on the manuscript have contributed sufficiently to be included as authors.

Thank you for the detailed reviews of our work and for the opportunity to re-submit.

Sincerely,

Ian D Graham, PhD
Reviewer: AM

p5. The three decision aids are referenced; however a brief sentence stating whether any have been demonstrated as effective, including refs, would be helpful.

A brief sentence describing the effectiveness of the three decision aids has been included along with the citations.

p6. Was there any further stratification of sampling (in addition to Dr-type), for example geographical location, age, postgraduate qualifications etc to ensure that representative samples were invited?

There was no further stratification. It was a random sample for both the family physicians and respirologists, and the entire eligible population for the geriatricians.

p6. The sample size calculation just needs a little clarification. I think by stating 150 per group, it implies separate analyses for each physician group, and n=150 would not allow a detectable margin of error of 10% unless the proportion with the outcome was at least 75%. Later it is clear that main analysis includes all three physician groups together.

We have modified the text to reflect that the sample size calculation was based on the total sample.

p6. I’m not sure I agree that a survey response rate of 60% is conservative, also borne out by the actual response!

The response rate was lower than anticipated. We address this limitation in the discussion section (pg.15).

p8. The decision to include respondents with 1 or 2 missing data items seems a sensible pragmatic approach – is there a reference for this approach? Also, imputing values based on the mode (or other summary measure of the observed data) can cause problems by underestimating the variance of a particular variable. Other methods, such multiple imputation using regression methods, are available and are usually produce more conservative (ie larger) variance estimates.

We do not have a reference for this approach. We agree that other methods for imputation could have been used.

p8. I am not an expert on factor or principal components analysis. However, aside from face validity, aren’t there usually other criteria (eg eigenvalues, cronbach’s alpha) used to determine the eventual number of factors/components?

Chronbach’s alpha was used for item analysis. The final four factors had Cronbach alphas ranging from 0.744 (Implementation – 6 items) to 0.921 (Quality & Value for Patients -12 items). The choice of four factors was also based on: 57% explained variance; similarity to hypothesized constructs in the scale; more than 4 factors produced factors with less than 3 factor loadings; scree plot; and parsimony.

p9. How were the components scored for the regression analyses vs intended use of DA?
Composite variable were created using the mean score of all 5-point Likert-scaled items within the factor. The text has been modified to reflect this point.

9. A little more information on the modelling strategy would be helpful here eg all univariable associations examined first, then variables with associations p<.2 entered into multivariable regression.

We have added further information to clarify the modeling strategy.

9 and table 1. The respondents are described – are there any similar national data available to enable comment on the generalisability of the sample?

Unfortunately we do not have information on the non responders.

9-11. I found much of this text repeated what was in the tables and became rather tedious. For me, the main analyses of interest are those in tables 5 & 6. Can the authors reduce the amount of text relating to the earlier analyses by just picking out one or two of the main points?

We have attempted to reduce some of the text on these pages. However, we feel it is important to highlight the differences in how respondents viewed each decision aid. Each physician group viewed a different decision aid and we have been careful to word the text to reflect this.

Results tables. Please give all p-values rather than designate those with p>0.05 as 'non-significant'.

We have adjusted the tables accordingly.

12. Terminology. Conventionally, univariate and multivariate are used to describe analyses with 1 and >1 outcome variables respectively; for multivariate, this means the simultaneous analysis of more than one outcome variable in the regression model. PCA is an example of a multivariate analysis. Univariable (which I think is what is actually meant by use of the term bivariate on line 18)

We have changed the text to reflect the above terminology

12. The penultimate sentence (“While physicians…”) is not a proper sentence.

We have made the appropriate grammatical correction.

12. The last sentence states that “Family physicians were three times more likely…” – I think the OR in table 5 to which this refers is actually 4.2.

We agree, the error has been corrected in the text.

Table 5. Please give N’s and denominators for each cell, rather than just ORs, 95% CI and p-values. eg create 2x2 table for comfort offering DA versus intention to use DA, etc. For predictor variables with >1 level, instead of giving p-values for each level compared
with reference category (analogous to sub-group analyses), should just quote single p-value from Wald or partial LR tests.

The suggested changes have been made in Table 5.

Discussion. As I mentioned in my general comments, I think the main limitation of this study is in the use of intended, rather than actual, behaviour as the outcome measure. Data on actual use were not collected from everyone. However might there be value in a secondary analysis that assumes non-use for everyone who did not say they used it at 3 months? The discussion focuses on self-report of behaviour rather than intention vs actual (whether self-reported or measured in some other way). As we all know, good intentions don't always translate into changes in behaviour, and I think this deserves much more attention in the discussion. I would guess that future work would look at predictors of actual behaviour, measured in ways other than just self-report, such as logged use via the web.

We feel that the sample size of those who indicated that they would use the decision aid is too small (n = 141) to warrant this analysis. We also feel that the analysis, as it stands, highlights the significant intention-behaviour gap that exists for an intervention that received overall positive support for its content and format. We agree that future work should examine predictors of actual behaviour.

Reviewer: MP

I might suggest that the proportion of physicians who adopted the decision aid should be viewed in more of a "half-full" rather than "half-empty" sense. I am surprised that even 32% would have actually implemented the decision aid. The QI literature is replete with similar levels of uptake for new innovations.

We agree that a 32% uptake could be viewed positively.

page 5 and discussion: I would not be quite as definitive about the essential role of physicians for the viability of patient decision aids. It is possible that decision aids could be directed to patients whether or not their physicians endorsed or used them. I am not ready to put all of our eggs in the basket of physician uptake.

We agree, we have tempered the tone of the text.

The authors should be careful in reaching conclusions about the different specialties, since the observed differences could either result from specialty difference or the difference in the clinical question being considered. Because all specialties did not comment on the same single decision aid, these effects cannot be separated.

We have tried to further address this issue in the limitations section of the discussion (pg 15).

The authors should report the proportion of those initially contacted to determine eligibility for the mail survey. (page 6, middle)
During the study, there were 3 research assistants and unfortunately, with the changes in personnel and the passing of time we have been unable to locate the files that would have contained this information.

The 32% figure is based on n=99 - it is probable that of those 141 who were eligible, the 42 that did not respond probably did not use the decision aid, either; thus 32% is an upper bound.

We agree this is likely. However, we chose to report data that we were able to obtain, while being explicit about the true numbers involved.

The authors should differentiate the challenges of performing shared decision making from the challenges of using decision aids. Decision aids were developed to address the challenges of performing SDM.

We agree that there is a differentiation between the challenges of performing shared decision making and challenges of implementing decision aids. However, there is some overlap of the barriers and facilitators to both. We have attempted to focus the discussion on implementation of decision aids in clinical practice.