Author's response to reviews

Title: Development and Psychometric Properties of a Brief Measure of Subjective Decision Quality for Breast Cancer Treatment

Authors:

Ken Resnicow (kres@umich.edu)
Paul Abrahamse (pabraham@med.umich.edu)
Rachel S. Tocco (rtocco@med.umich.edu)
Sarah Hawley (sarahawl@med.umich.edu)
Jennifer Griggs (jengrigg@med.umich.edu)
Nancy Janz (nkjanz@umich.edu)
Angela Fagerlin (fagerlin@med.umich.edu)
Adrienne Wilson (deebz@umich.edu)
Kevin C. Ward (kward@emory.edu)
Sheryl G.A. Gabram (sgabram@emory.edu)
Steven Katz (skatz@med.umich.edu)

Version: 2 Date: 20 October 2014

Author's response to reviews: see over
RESPONSE TO REVIEWERS
MS: 7635101391379463

Development and Psychometric Properties of a Brief Measure of Subjective Decision Quality for Breast Cancer Treatment

Title: Development and Psychometric Properties of a Brief Measure of Subjective Decision Quality for Breast Cancer Treatment

Reviewer: IS

Comment: Many thanks for the opportunity to review this paper. It is a well written manuscript. Short measures are needed in the field of shared decision making and this paper adds to the currently available measures. However, I have a range of comments and suggestions to strengthen the manuscript.

Response: Appreciated

Comment 1: Please adhere to the COSMIN checklist for reporting your results.

Response: We reviewed the COSMIN checklist and the associated coding manual from www.cosmin.nl. We have completed the form, of course using the CTT rather than IRT criteria. Rather than strictly using their checklist to structure the paper, we have made sure that all of the relevant criteria are addressed, and when not, their absence is listed as a limitation. We are happy to indicate that the completed form can be obtained from the authors. We have attached a completed version for you to review and modify (in the event you have a different take on our psychometrics).

Comment: 2. What the authors define as “subjective decision quality” (Introduction, p.4) has often been described in the literature as decision process and decision outcomes. Many measures of those aspects exist (e.g. Decisional Conflict Scale, Satisfaction with Decision Scale, Decision Regret Scale, SDM-Q-9 and many more). The manuscript lacks any acknowledgement and discussion of this prior work. The authors should be more clear why a new measure is needed, i.e. that there is a lack of short measures. It also remains unclear why convergent validity was not assessed by the use of these available measures. This should at least me noted in the discussion section.

Response: We have now included the distinction between decision process and decision outcomes. And, we also now refer to several prior measures, and have better justified the need for a new instrument, and our desire to develop a brief measure that could be used globally for preference sensitive decisions. Specifically, we have added the following text:

Although the distinction between objective and subjective elements of breast cancer or medical decision making more broadly, has not been previously proposed, this dichotomy mirrors the distinction between the deliberation process and the determination of decisions proposed by Elwyn and Miron-Shatz 25. Specifically the authors note (ibid page 143), “we need to evaluate both the decision making (the perceived or observed process) and the determination (whether or not the decision itself is considered “good””). Our proposed measure similarly distinguishes between the objective processes used in making the decision and the subsequent subjective evaluation of that decision, with a focus on the latter.
And we added,

Numerous objective knowledge and concordance measures, particularly for the purpose of evaluating patient decision aids, have been reported, including the Decisional Conflict Scale and some consensus regarding the content of these measures has emerged. There are, however, fewer measures of subjective decision quality and even less consensus regarding how the construct should be assessed. We could find only one measure that exclusively tapped subjective decisional quality: the Satisfaction With Decision (SWD) scale by Holmes-Rovner et al. This measure however, only assesses positive aspects of Decision Quality, i.e., it contains no negatively framed items such as regret and does not assess satisfaction with how much involvement and time the patient perceives they had in making their decision, both of which have been proposed as key elements of a quality decision. The proposed measure evaluated herein, includes these elements.

Although we did not delve into this degree of detail in the paper, we did want to explain in more detail our thoughts on a few of the prior measures. Let’s start with the Decisional Conflict Scale. As the authors of the scale indicate, the 16 items form five factors, 1) Informed, 2) Values Clarify, 3) Support 4) Uncertainty, and 5) Effective Decisions.

### 16 Decisional Conflict Scale (DCS) items

1. I know which options are available to me.
2. I know the benefits of each option.
3. I know the risks and side effects of each option.
4. I am clear about which benefits matter most to me.
5. I am clear about which risks and side effects matter most.
6. I am clear about which is more important to me (the benefits or the risks and side effects).
7. I have enough support from others to make a choice.
8. I am choosing without pressure from others.
9. I have enough advice to make a choice.
10. I am clear about the best choice for me.
11. I feel sure about what to choose.
12. This decision is easy for me to make.
13. I feel I have made an informed choice.
14. My decision shows what is important to me.
15. I expect to stick with my decision.
16. I am satisfied with my decision.

The Informed Subscale items, which comprise I know which options are available to me, I know the benefits of each option, and I know the risks and side effects of each option, we feel more closely map onto objective process dimension. And of course our scale focuses on the subjective outcome evaluation. The DCS does not assess how satisfied or adequate the patient feels about their level of knowledge. Similarly, the Values Subscale items, I am clear about
which benefits matter most to me, and I am clear about which risks and side effects matter most, appear to tap awareness rather than satisfaction. The Support Subscale, I have enough support from others to make a choice, I am choosing without pressure from others, and I have enough advice to make a choice taps adequacy of support, which we conceived as a separate dimension or even predictor of subjective decision quality. The Uncertainty Subscale, comprised of, I am clear about the best choice from me, I feel sure about what to choose, and This decision is easy for me to make combines several elements including confidence and ease of the decision, neither of which are the focus of our SDQ. Finally, the Effective Decision Subscale, comprised of I feel I have made an informed choice, My decision shows what is important to me, I expect to stick with my decision, and I am satisfied with my decision includes some elements of our SDQ but includes future intention. Thus, we feel there is only minor overlap between the DCS and our proposed measure. We do feel it would be useful to compare values on the two, and we have highlighted this in our discussion section.

We also examined the Satisfaction with Decision Scale by Holmes-Rovner et al, as this scale appears to be more closely related to our proposed scale. Their scale comprises six items:

1. I am satisfied that I am adequately informed about the issues important to my decision.
2. The decision I made was the best decision possible for me personally.
3. I am satisfied that my decision was consistent with my personal values.
4. I expect to successfully carry out (or continue to carry out) the decision I made.
5. I am satisfied that this was my decision to make.
6. I am satisfied with my decision.

We feel it taps global satisfaction (items 2 and 6) and satisfaction with their degree of information (item 1), both of which are included on our scale. Item 3, which taps values concordance, we chose not to include as we found it somewhat difficult for lower literacy patients to comprehend. This SWD measure does not tap the dimensions of having perceived adequate time or involvement. And, we feel that by adding future intention (item 4) they have gone beyond what we are calling subjective decision quality.

Finally, apropos to both the DCS and SWD measures, we previously found that the response categories they use, strongly disagree to strongly agree can evoke higher rates of acquiescence and endorsement of the extreme positive end of the response scale. In our current SDQ measure, as well as prior measures from our group, we have moved to a different response range, i.e., NOT ENOUGH, JUST RIGHT, and TOO MUCH, so that the criterion response lies in the middle of the response continuum rather than on one of its poles.

Comment 3. A main flaw of the current manuscript is the lack of description of scale development. The title suggests that the manuscript will report on the development of the tool, but this is not mentioned in the aims of the study nor is the scale development adequately described. However, this is extremely important to be able to appraise the quality of the measure in terms of content validity (see COSMIN checklist or textbooks on scale development). How where the items created exactly? Did the authors involve patients in the scale development? Where pre-tests (cognitive debriefing, etc) done? Etc

Response: Most of the items we used in this scale were already pretested in prior studies, although not necessarily in the same scale. Thus, we did not specifically pretest them for this psychometric pilot. We realize this would have been helpful and we now note that lack of pretesting, specifically cognitive testing is a weakness of our scale development process. We have acknowledged this in the discussion with the following text,
“Finally, we did not conduct cognitive interviews with patients for this psychometric pilot. Although we had used these items and the corresponding response scales in prior studies and obtained respondent feedback, it would have been useful to conduct further pretesting including cognitive interviewing here to better elucidate how the items are perceived. This is a priority for our future research.”

Comment 4. It remains unclear what the item “right for you” means. Has this been checked beforehand in pre-tests? The results also indicate that it might have been difficult to understand this item.

Response: We concur that this item may not have been well understood. Above we note our lack of cognitive interviewing as a limitation. Moreover, based on the factor loading values we have decided to recommend that it should be removed from the scale, leaving 5 rather than 6 items.

Comment 5. Please discuss the ceiling effects.

Response: This is an important issue. We have now added the following text in the results,

....Across the six items between 70% and 89% of respondents answered with the most positive category, i.e., completely Right for Me, No Regret, or Totally Satisfied for the right for me, regret, and satisfaction items, respectively, and Just Right for the information, involvement, and time items.

And we added the following text in the discussion;

“The overall means of the four scales were high; 6.3 or higher (out of 7) and across the six items between 70% and 89% of respondents answered with the most positive category, i.e., completely Right for Me, No Regret, or Totally Satisfied for the right for me, regret, and satisfaction items respectively, and “Just Right” for the information, involvement, and time items. ...................... Additionally strategies to reduce positive respondent bias, perhaps by altering our item stems or response categories merits attention.”

Comment 6. In the conclusion you should clearly point out the need for further psychometric testing, both convergent and divergent validity, as well as sensitivity to change and other aspects of psychometrics.

Response: We concur and have added the following text,

“Future psychometric testing is encouraged, including establishing norms for the scale as well as its convergent and divergent validity by examining how the measure is associated with other outcomes such as quality of life, decisional conflict, psychological coping, objective decision quality, and perhaps even survival.”

Minor Essential Revisions

Comment 7. Please provide the developed questionnaire as a table, rather than just plain text.
Response: We have included the exact items as an Appendix.

Comment 8. Please add keywords related to scale development and psychometric testing.

Response: Done

Comment 9. The abstract needs revision, especially its conclusions. I struggle with the sentence “Appraisal of breast cancer treatment decisions comprises six dimensions”, as you found only one factor in your analyses. This should also be revised in the discussion. Also, I think that you should add to the conclusion that further psychometric testing of the measure is necessary.

Response: We have rewritten the conclusion of the abstract which now reads;

“Our measure of subjective appraisal of breast cancer treatment decisions includes related elements; regret and satisfaction as well as perceived adequacy of information, time, and involvement. Future research is needed to establish norms for the measure as is further psychometric testing, particularly to examine how it is associated with outcomes such as quality of life, psychological coping and objective decision quality.”

We have made similar changes to the discussion/conclusion section which now reads;

“Future psychometric testing is encouraged, including establishing norms for the scale as well as its validity by examining how the measure is associated with other outcomes such as quality of life, psychological coping, objective decision quality, and perhaps even survival.”

Comment 10. Methods: Please explain the dichotomization of demographic characteristics.

Response: For age and months since diagnosis, we dichotomized the sample around the median. We felt this allowed for adequate sample sizes to make comparisons between groups, and presentation of our results would be simplified by using two strata for each variable rather than providing a regression coefficient for a continuously scaled variable.

Comment 11. Methods: How long where the decisions ago? Couldn’t there be a recall bias?

Response: We concur that recall bias could influence our results. We capped the maximum time since diagnosis allowed into the study at 18 months. As shown in Table 4, mean scores on the four scales did not differ by groups stratified by time since diagnosis (split at the median of 12 months). Below are the data directly pulled from Tab 4.

<table>
<thead>
<tr>
<th>Months since Diagnosis</th>
<th>&lt; 12</th>
<th>131</th>
<th>6.52</th>
<th>130</th>
<th>6.61</th>
<th>83</th>
<th>6.35</th>
<th>114</th>
<th>6.65</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&gt; =12</td>
<td>167</td>
<td>6.39</td>
<td>165</td>
<td>6.51</td>
<td>124</td>
<td>6.32</td>
<td>141</td>
<td>6.48</td>
</tr>
</tbody>
</table>

Comment 12. Are the decisions to undergo surgery, chemo and radiation really independent decisions? I would doubt this from a clinical perspective. Does it than make sense to ask to rate them separately? Could the association of scores be due to association of treatments?
Response: You have posed a critical question for which there is no simple answer. Ultimately it is a “glass half full/half empty” issue. As we note in the manuscript,

“The three treatment-specific decision scale scores, computed from the means of the six items, formed a single factor with an eigenvalue of 1.75. The factor loadings for the three treatment-specific subscales on this single factor were .75, .76 and .79, for surgery, chemotherapy, and radiation, respectively. The correlations of the three subscale scores of decision quality with the overall decision quality score were .75, .70, and .74, respectively, for surgery, chemotherapy, and radiation.'

Later in the discussion we note,

“A second finding is that scores from the three treatment-specific decisions scales correlated with overall decision quality in the range of .70 to .75, with surgery exhibiting the highest correlation with the overall score. This raises the question as to whether decision quality should be assessed for each treatment separately or only for the patient’s overall treatment experience. On one hand, given these correlations, each treatment appears to capture a somewhat independent picture of decision quality. However if questionnaire space is limited and/or respondent burden is a particular concern, measuring only the overall treatment experience may be appropriate.”

On one hand, given that the correlations of the three subscales (surgery, chemotherapy, and radiation) with overall DQ, are in the range of .70 to .75, it could be argued that they are at least moderately independent aspects of overall DQ. Additionally, the fact that the three subscales load about equally on a single underlying dimension could also be interpreted as a need to measure all three dimensions. If the factor loading and scale to total score correlations were demonstrably higher for surgery for example, then we might conclude that surgical experiences contributes more to the total score, and therefore assessment of the other domains would not be warranted. However, this is not what we found. The fact that correlations of the subscale scores with total DQ were all around 0.7, suggest that we capture a large percentage of the variance in the subscales just by using the DQ overall score. This is why we recommend that,

“..., if questionnaire space is limited and/or respondent burden is a particular concern, measuring only the overall treatment experience may be appropriate.”

Comment 13. In the Discussion section, the authors describe an “ordering effect” (p.12). Please report this in the Results section.

Response: Given the fact that we are using the ordering results to explain our findings and they are only tangential to the primary psychometric results, we prefer to leave this section in the discussion. If the reviewers and editors feel otherwise we can easily move this section.

Comment 14. Is “time since diagnosis” the right indicator to assess whether scales vary or could it also be time since decision or time since treatment? (p.12.)

Response: Unfortunately we cannot answer this empirically, as we did not measure time since treatment or time since decision, only time since diagnosis. We did this in part because we have found that patients are better able to recall their diagnosis date than dates of their treatment decision or treatment initiation.
Comment 15. The term “objective” (Introduction, p.3, line 42) is often associated with observer measures. You could clarify that you talk about patient-reported measures to avoid confusion.

Response: This distinction is captured in our new text that cites Elwyn and Miron-Shatz,

Although the distinction between objective and subjective elements of breast cancer or medical decision making more broadly, has not been previously proposed, this dichotomy mirrors the distinction between the deliberation process and the determination of decisions proposed by Elwyn and Miron-Shatz \(^{25}\). Specifically the authors note (ibid page 143), “we need to evaluate both the decision making (the perceived or observed process) and the determination (whether or not the decision itself is considered “good”)

Our proposed measure similarly distinguishes between the objective processes used in making the decision and the subsequent subjective evaluation of that decision, with a focus on the latter.
Title: Development and Psychometric Properties of a Brief Measure of Subjective Decision Quality for Breast Cancer Treatment

Reviewer: KS

Reviewer's report:
The authors report on initial psychometric analysis of a new survey instrument to measure the quality of decisions for breast cancer treatment. The survey includes six items each representing a different dimension including regret, satisfaction, fit, adequacy of information, and involvement. They examined the items with 320 women who reported on their decisions about surgery, chemotherapy and radiation. It is a well written manuscript describing an initial field test of the instrument. Although there are several instruments available that assess different aspects included in the subjective measure tested here, having a brief one that adequately captures the different elements could be helpful. However, there are several areas of the manuscript would benefit from additional details and evidence.

Major revisions:

1. Perhaps most important would be a clearer description of the theory or framework used to define decision quality. The authors describe six dimensions, but do not indicate why these six were used as opposed to others. Formal decision analysis has six dimensions of decision quality (but not these six), the IPDAS standards and others have defined decision quality with three dimensions (informed, involved, value concordant), etc. Please include more details on the rationale, theory and or framework used to guide the selection of these dimensions (and why other dimensions were not included).

Response: Good point. We have added the following text to describe our theoretical orientation.

“Conceptually, our measure was also informed by Self-Determination Theory (SDT), which distinguishes between controlled and autonomous motivation. SDT posits that behavior change decisions that respect and support the patient’s autonomy and competence are more likely to be perceived as higher quality than decisions that leave the patient feeling pressured or compliant (i.e., controlled motivation). Specifically, the item regarding adequate knowledge is intended to tap competence, whereas items related to involvement and time are intended to tap the degree of autonomy in the decision making process.”

In addition we added,

“Although the distinction between objective and subjective elements of breast cancer or medical decision making more broadly, has not been previously proposed, this dichotomy mirrors the distinction between the deliberation process and the determination of decisions proposed by Elwyn and Miron-Shatz. Specifically the authors note (ibid page 143), “we need to evaluate both the decision making (the perceived or observed process) and the determination (whether or not the decision itself is considered “good”). Our proposed measure similarly distinguishes between the objective processes used in making the decision and the subsequent subjective evaluation of that decision, with a focus on the later.”
We also have examined the IPDAS criterion for measuring patient decision aids and now cite this important work (Sepucha KR, Borkhoff CM, Lally J, Levin CA, Matlock DD, Ng CJ, et al. Establishing the effectiveness of patient decision aids: key constructs and measurement instruments. BMC Med Inform Decis Mak. 2013;13 Suppl 2:S12). One key difference between our measure and the IPDAS standards is that the latter were created to guide the evaluation of decision aids, whereas we are more concerned with patient decisions primarily as they relate to interaction with their physicians. We would nonetheless like to provide some detail about how our measure fits into the IPDAS framework. There appears to be two domains Sepucha et al recommend for evaluating decision aids, 1) the quality of the decision-making process and, 2) the quality of the choice. Sepucha et al note that dimension 1 comprises several subconstructs:

1. Recognize that a decision needs to be made (e.g., as measured by items in the Preparation for Decision Making Scale (PMDS) [3].
2. Feel informed about the options and about the risks, benefits, and consequences of the options (e.g., as measured by the “Feeling Uninformed” subscale of the Decisional Conflict Scale [4]).
3. Be clear about what matters most to them for this decision (e.g., as measured by the “Unclear Values” sub-scale of the Decisional Conflict Scale (DCS) [4]).
4. Discuss goals, concerns, and preferences with their health care providers (e.g., as measured by items in the Perceived Involvement in Care Scale (PICS) [5]).
5. Be involved in decision making (e.g., as measured by the Control Preferences Scale (CPS) [6] and adaptations of the CPS).

In general our measure is more aligned with quality of the choice rather than the decision process. Specifically, constructs 2 and 5 above, we capture with our items on Informed and Involved. However, we are focusing on the subjective rating of these elements rather than the objective process. Constructs 1 and 3 seem to relate more to processes/outcomes specific to decision aids, and may be less relevant to our measure which focusses on the provider-patient aspect of the decision rather than interaction with a decision aid tool. Construct 4 relates to provider communication, which we partially tap in our measure with the Involvement item. However, in addition, we usually include the HCCQ (Health Care Climate Questionnaire by Williams et al) in our questionnaires as a separate measure of patient provider communication. In a future publication we will report the association between our SDQ measure and the HCCQ.

Dimension 2 (i.e., quality of the choice) comprises two sub-elements, 1) how informed, and 2) how concordant. We feel using our taxonomy these both map onto “objective” measures of outcomes rather than process. Our measure differs in that we are primarily assessing patients’ subjective satisfaction with their degree of being informed and how well their physician involved them in the process, rather than the objective nature of the process. This is an important distinction because a patient may not be fully informed about all options but may nonetheless feel they had enough information or they may be fully informed but feel it was too much to process and may report low satisfaction. It is not entirely clear from the IPDAS criteria whether these constructs are recommend to be measured as objective or subjective elements. We believe it is a potential benefit of our scale that it clearly distinguishes between the objective and subjective. As noted in the discussion, how our subjective measure, might be associated with more objective measures such as actual knowledge level and values concordance is a high priority for future research.

**Comment 2.** Further, it would be helpful to understand more about the item development. I was surprised to see only one item used for each dimension as many of these dimensions have entire surveys devoted to them e.g. Decision Regret Scale, Decision Making Satisfaction Scale,
and Decisional Conflict Scale, etc. Please include more details on how the wording of the items were determined, why there is only one item for each dimension, and any cognitive testing or other work to ensure respondents understand the items as intended.

Response: Our goal was to develop a brief measure, thus leading to our decision to have one item per dimension. We feel the alpha levels obtained support the decision to use a single item per construct.

Comment 3. It would be helpful in the introduction and again in the discussion to present a more compelling case for the need for such a survey. What gap is this filling? How are these items providing better, new, more relevant information than current scales for decision quality (see examples listed in prior comment)?

Response: We concur and have added the following text.

“Although the distinction between objective and subjective elements of breast cancer or medical decision making more broadly, has not been previously proposed, this dichotomy mirrors the distinction between the deliberation process and the determination of decisions proposed by Elwyn and Miron-Shatz. Specifically the authors note (ibid page 143), “we need to evaluate both the decision making (the perceived or observed process) and the determination (whether or not the decision itself is considered “good”). Our proposed measure similarly distinguishes between the objective processes used in making the decision and the subsequent subjective evaluation of that decision, with a focus on the latter.”

In addition, we added the following,

“…There are, however, fewer measures of subjective decision quality and even less consensus regarding how the construct should be assessed. We could find only one measure that exclusively tapped subjective decisional quality; the Satisfaction With Decision (SWD) scale by Holmes-Rovner et al. This measure however, only assesses positive aspects of DQ, i.e., it contains no negatively framed items such as regret and does not assess satisfaction with how much involvement and time the patient perceives they had in making their decision. The proposed measure evaluated herein, includes these elements.”

Comment 4. The lack of any strong framework makes some of the results difficult to interpret. For example, with Table 4, did the authors have any hypotheses about how these items would work across the different decisions? Across different patients?

Response: We did not have a priori assumptions above how different patient populations would respond to these items, although we feel these results, even if atheoretical, are nonetheless worth reporting.

Comment 5. Despite finding that only five dimensions appeared to hang together psychometrically the authors recommend keeping all six. Why? Did they examine whether it was an issue perhaps with wording of the item? Or does that dimension need more than one item to assess adequately? A little more detailed discussion of this result and implications for future versions of the survey would be helpful.
Response: You make an excellent point. We have revised our recommendation regarding the inclusion of the Right for You item. We now recommend the 5-item version, which excludes this item. Here is the new text.

“.The primary finding from these analyses is that subjective appraisal of breast cancer treatment decisions appears to comprise at least five related dimensions. These five dimensions, regret and satisfaction as well as adequacy of information, time, and involvement appear to merit inclusion on the final scale for the overall treatment as well as treatment-specific scales (i.e., surgery, radiation, and chemotherapy). The remaining element “Right for You”, was a candidate for removal, based on factor loading and/or internal consistency results for decisions regarding chemotherapy and radiation. For overall treatment and surgery specific decisions, however, it appears to merit inclusion. Although for parsimony we recommend the five –item version, for investigators assessing decision quality for chemotherapy and radiation, inclusion of the “right for you” item may be justified.”

Minor revisions:

Comment 6. The authors present this as measure of breast cancer decision quality – however the dimensions seem fairly generic. Was there any specific adaptation or work to make these relevant for breast cancer? Were breast cancer patients and providers involved in identifying the dimensions or creating the items?

Response: We were not clear on this issue. Our aim was to develop a measure that could be used for other treatment decisions beyond breast cancer, particularly other preference sensitive decisions. We have added this point to intro and the discussion. Specifically we now state in the introduction,

“The primary aim of this study was to explore the psychometric properties of a new subjective quality decision measure for breast cancer treatment that could be applied to other preference sensitive decisions.”

And, in the conclusion,

“Although we tested the measure in the context of breast cancer treatment, the scale may be applicable to assessing subjective decisional quality for other preference sensitive medical decisions.”

Comment 7. Please provide some details on the selection of the 7-point response scale and on the performance of the items. What was the range of responses for each item (e.g. how often were each of the 7 responses used for each item)? Please describe in more detail.

Response: As noted above there was score compression on the upper end. This is reflected in the high overall means (around 6.5) and overall around 80% of respondents answered a 7. We agree this is a limitation of the scale. To address this issue we added the following text in the results:

….Across the six items between 70% and 89% of respondents answered with the most positive category, i.e., completely right for me, no regret, or totally satisfied for the right
for me, regret, and satisfaction items respectively, and Just Right for the information, involvement, and time items.

And the following in the discussion;

“The overall means of the four scales were high; 6.3 or higher (out of 7) and across the six items between 70% and 89% of respondents answered with the most positive category, i.e., completely right for me, no regret, and totally satisfied for the right for me, regret, and satisfaction items respectively, and Just Right for the information, involvement, and time items. ……………… Additionally strategies to reduce positive respondent bias, perhaps by altering our item stems or response categories merits attention.”

Comment 8. Please provide more details on response rates and missing data, especially as it pertains to understanding the acceptability and feasibility of the survey instrument (e.g. only 195/320 had all items complete for chemotherapy). Which items were skipped? It would also be helpful to clarify the survey protocol e.g. whether reminders were used for patients who took the survey home to complete. Please discuss in the results and limitations the implications of such a large number of incomplete surveys and missing data.

Response: We now provide data about missing and skipped items. We added the following text,

“The number endorsing “not offered” across the six items for overall treatment ranged from 2 to 8, for surgery, 7 to 12, for radiation, 47 to 57, and for chemotherapy, 83 to 102. The skip rate across the six items for overall treatment ranged from 2 to 8, for surgery 2 to 6, for radiation, 5 to 12, and for chemotherapy 3 to 6 respondents.”

In terms of reminders, we added the following text to clarify our procedures,

“To increase response, we mailed reminder postcards, made 1-2 follow-up calls and re-mailed a second packet (with no second incentive). “

Comment 9. Please comment on ceiling effects and whether there is a need to revise the items in any way to remove those.

Response: Please see our response to your comment #7 above.

Comment 10. It seems premature in the conclusion to recommend assessing these six dimensions of decision quality when the paper did not present any other dimensions (or compelling rationale as to why these six were particularly important) nor did the authors present any compelling evidence of validity (except to argue that patients’ perceptions are inherently valid).

Response: We have tempered our conclusion in this regard. Here is the new relevant text.

“. The primary finding from these analyses is that subjective appraisal of breast cancer treatment decisions appears to comprise at least five related dimensions. These five dimensions, regret and satisfaction as well as adequacy of information, time, and involvement appear to merit inclusion on the final scale for the overall treatment as well as treatment-specific scales (i.e., surgery, radiation, and chemotherapy). The remaining element “Right for You”, was a candidate for removal, based on factor loading and/or
internal consistency results for decisions regarding chemotherapy and radiation. For overall treatment and surgery specific decisions, however, it appears to merit inclusion. Although for parsimony we recommend the five–item version, for investigators assessing decision quality for chemotherapy and radiation, inclusion of the “right for you” item may be justified.”

More expansive exploration regarding the validity of our decision quality measure, however, is warranted. This could include examining how our measure may be associated with outcomes such as quality of life, cancer-related anxiety, and perseveration and personality attributes such as social desirability and optimism. ….Our survey did not include objective measures of decision quality, i.e., treatment knowledge, values concordance, or decisional concordance. Examining the relationship between objective and subjective measures of decisional quality is encouraged.

Future psychometric testing is encouraged, including establishing norms for the scale as well as its validity by examining how the measure is associated with other outcomes such as quality of life, psychological coping, objective decision quality, and perhaps even survival. Although we tested the measure in the context of breast cancer treatment, the scale may be applicable to assessing subjective decisional quality for other preference sensitive medical decisions.