Author's response to reviews

Title: Values Clarification in a decision aid about fertility preservation: does it add to information provision? Two randomized experiments with healthy participants

Authors:

Mirjam M Garvelink (m.m.garvelink@lumc.nl)
Moniek M ter Kuile (m.m.ter_kuile@lumc.nl)
Anne M Stiggelbout (a.m.stiggelbout@lumc.nl)
Marieke de Vries (mariekedevries@uvt.nl)

Version: 3 Date: 11 April 2014

Author's response to reviews: see over
Dear Dr Dolan,

We would like to thank the editor for the kind opportunity to revise and resubmit our manuscript "Values Clarification in a decision aid about fertility preservation: does it add to information provision? Two randomized experiments with healthy participants (MS: 1848265293117094)" Herewith we resubmit a revised version of our manuscript for review to BMC Medical Informatics and Decision Making.

In the following we will describe point by point how we have responded the comments of the reviewers. All reviewer comments were taken into account. For the purpose of clarification, we have indicated in the revised paper where changes have been made, and we have keyed these to the reviewers’ comments below. Attached you will find the revised manuscript.

We thank you and the reviewers for your very constructive comments that have helped to further improve our manuscript. We hope the manuscript in its present form will be now found acceptable for publication in BMC medical informatics and decision making.

Yours sincerely, on behalf of my co-authors,
Mirjam Garvelink

Reviewer 1

1. It is difficult to understand what VCE++ means in the methods section of the abstract. Since the abstract should makes sense as a stand-alone piece can the authors include a definition of VCE++ please.
   As was requested by the reviewer we have included a more detailed definition of VCE++ to the abstract of the manuscript. We have added (in red): “information plus a VCE with specific referral to the VCE, requesting participants to use the VCE (VCE++)”

2. There is a new update just published for the Cochrane systematic review of PtDAs (ref 3) which would be better for the readers of the journal to have as a reference.
   As was requested by the reviewer we have updated our third reference with the updated Cochrane review on PtDAs. Thank you for notifying us.

3. Can the authors provide some explanation of what was contained in the different PtDA versions? What was the information content? What type of VCE was used? How were participants referred to the VCE in Exp 2 in the VCE++ arm? Other researchers who may wish to build on this work should be able to see what has been done. In other words, there needs to be a better description of the intervention in each arm of the studies.
In response to reviewer 1 we have provided more details about the differences between the two different versions of the DA, and about the procedure of referral to the VCE in experiment 2. We have added a figure (figure 3) that visualizes the different versions of the DA. Also, in the text (experiment 1 and 2, methods section, procedure) we have added: “[..] using the following text: ‘by clicking on the link below you are referred to a decision aid about fertility preservation for breast cancer patients. You are asked to make a decision whether or not you want to preserve your fertility, and if so, how’.”

and we have additionally added to the method section of experiment 2: “Respondents who were specifically referred to the VCE additionally received the following text: “we would like to point out that the decision aid consists of both textual information about fertility preservation as well as the chapter “deciding about fertility preservation” which is meant to help you order your thoughts about fertility preservation and make a decision. Please use this chapter in making your own decision about fertility preservation”."

With regard to additional details about the content of the DA and VCE and considerations in development of it, see also reviewer 2, point 1.

4. Perhaps it could be noted that the VCE did not appear to cause any harm to participants other than the time involved in completing (which was not onerous). This is an important point given the recommendation to test this on 'real patients'.

As was suggested by reviewer 1 we have added the following to the conclusion section of the manuscript: “However, nor did use of the VCE seem to cause any harm, other than the time involved in completing it (which was acceptable).”
1. More information needs to be provided regarding the nature of the values clarification exercise.
Although the general format of the clarification procedure used in the study is outlined on page 7, there is not sufficient detail provided to fully understand what the exercise entailed, why it was expected to work, and it was implemented during the study. There also needs to be information provided regarding the theoretical basis, if any, of the clarification method used and if it has been shown to be effective in other contexts or pilot studies.

In response to both reviewers we have provided more details about the DA and VCE. We have added (in red): “In order to assist decision making about FP, we have developed a DA for women with breast cancer who have to decide about FP treatments [18]. The DA consists of textual information, and a fine-grained, explicit VCE. The few studies that have evaluated the impact of VCEs found indications for a beneficial effect of adding a VCE to a DA with regard to the percentage of patients who made an informed decision that was in agreement with their personal values [3; 14]. We have chosen for an explicit VCE since explicit VCEs showed indications of being more effective than implicit VCEs with regard to satisfaction with preparation for decision making [13] and decisional conflict [7]. We have chosen a fine-grained, additive VCE (comparing attributes of one treatment option at a time) [9], as we wanted patients to choose only between pursuing (or not) the options for which they are eligible. Each VCE consists of statements about the consequences of an FP option, for each of which patients were asked to indicate the extent to which these statements were considered a benefit or disadvantage (Figure 1,2). Based on considerations mentioned by Feldman-Stewart et al. (2006) [9], assuming understanding of a continuous scale underlying the bars in their VCE by a majority of respondents, and that bar settings were congruent with the chosen treatment option, we hereby chose two bars (with underlying VAS scales) to indicate a) whether the statement is considered to be an advantage or disadvantage to the FP option, and b) the importance of a statement [9] (figure 1). Additionally, patients have the option to add arguments and rate these as well. After rating the importance of the separate statements, the DA generates a summary that provides an overview of patients’ answers in descending order from most important to least important (as indicated by the patient). Moreover, patients can indicate the extent to which they favor the treatment options, and make a decision based on their own values. Patients are not provided with a clear-cut advice about which treatment to choose. This overview was chosen, rather than a summary bar indicating how much someone favors one of the FP treatments [9], because we did not want to steer patients towards one of the treatments. In previous studies with this DA, acceptability, comprehensibility and user-friendliness were assessed in patients and clinicians and both the textual information and the VCE were considered relevant, coherent and understandable [19]. We hypothesized that the use of our DA with VCE in
deciding about FP would decrease decisional conflict compared to information only [7;13].“

2. More details should be provided regarding the choice of personality and coping styles were chosen and what the constructs mean. These terms are used but never really defined so are therefore difficult to fully understand.

As was requested by reviewer 2 we have provided more details about the personality and information seeking styles that we have measured in experiment 2. In the methods section of experiment 2 (section “measurements”) we have added between brackets the styles that we have measured, and as explanations we have added to information seeking styles: “(monitoring and blunting) A monitoring information seeking style indicates cognitive confrontation; a person with this style tends to actively seek out and monitor information about the threatening event [32]. A blunting information seeking style indicates cognitive avoidance; a person with this style tends to seek cognitive distraction from the threatening event and psychologically blunts threat-relevant information [32]. “and to Personality traits: (neuroticism and conscientiousness) “A high score on neuroticism indicates that women are emotionally instable; a high score on conscientiousness indicates that women are well-organized and task- and goal-directed [34].”

3. More information needs to be provided about the decision aid.

Readers are told that the basic decision aid was long (20 and 26 pages for the version with and without the clarification exercise respectively) and originally designed for use in consultation with a clinician rather than for a decision by an individual patient. This information should be included in the methods section and a supplemental file containing the decision aid would be helpful. In addition, however, there should be additional discussion, beyond being simply noted as a limitation, of how this discrepancy between the way the decision aid was designed and how it was used for the study might affect the results. In response to the request by reviewer 2 to include information about the original purpose of the DA we have added to the introduction section: “The DA has been designed for use by individual patients, as preparation for a consultation with a clinician in which the final decision is made.” We have chosen for slightly different wording than suggested by the reviewer since the DA was not designed for use IN the consultation, but as preparation to the consultation. We have chosen to include this information in the introduction instead of the method section (as was requested by the reviewer) to keep all information about the DA, its purpose and its content together instead of spread throughout the manuscript. We also added mentioning of the discrepancy in the original design of the DA and how it was used in these experiments to the discussion instead of only mentioning it as a limitation (see comment 5, reviewer 2).

4. The data reported in Table 1 and the conclusion to experiment 1 is confusing. The sentence on page 15 “Secondary analyses within women who received a DA with VCE revealed less decisional conflict for women who used the VCE
compared to those who did not use it, but with no certainty that it was the VCE that caused this difference, since there was no difference when VCE-users were compared to women who received a DA with information only (without the VCE).” Does not seem consistent with the results presented in the Table. The post-hoc notation used in the table is quite difficult to follow. Please clarify.

In response to the reviewers comment that the conclusion of experiment 1 and the data in table 1 are confusing we have changed the formulation of the conclusion to: “Secondary analyses revealed less DC for women who used the VCE compared to those who chose not to use it. Since there was no difference between the VCE-users and the women who received a DA with information only (without VCE) this is likely an effect of VCE-use in a self-selected group, and not of the VCE per se.” We have also adapted the post hoc notation used in table 1 to make it more easy to follow. The notation is now similar to that in table 2.

5. An additional limitation not highlighted is whether these results can be generalized outside of the study context.
More discussion has to be included regarding the extent to which the results of a study done with participants who were not facing a real decision, using long, text-heavy decision aid originally designed for use in consultation with a provider, studying a values clarification exercise that has not been shown to be effective (at least based on the information provided in this manuscript) can be extrapolated to other decision aids and settings.
Reviewer 2 requested that we highlighted whether results can be generalized outside the study context. Since the primary limitation mentioned is generalizability we have not highlighted it additionally in the limitation section, but we did mention it again in the paragraph of the discussion section where we discuss the differences between studies with healthy participants and patients. We have added: “This may also be related to discrepancies between the way DAs are designed and how they are used in healthy participants. It should be noted that the DA as used in the experiments was originally designed for patients, who use the DA in preparation for a consultation with a physician in which a final decision is made about FP. This consultation is often within a few days after diagnosis (and DA use). In the experiments, respondents had to decide directly after viewing the DA, without support from a physician. Hence, both the limited amount of available time and the lack of interaction about the decision may have influenced decision making for our participants. It is likely that in the experiments decisions were made consciously since they were made directly after viewing the DA. Actual patients may make more intuitive decisions, since they are distracted in the time between using the DA and visiting the physician to decide. Sometimes, decision making may improve when the decision is made after distraction, due to the so-called unconscious thought effect [11,50].

6. The lack of any effect of the emotional induction in experiment 1 is a noteworthy finding that should be discussed, if only briefly.
In response to the reviewers comment we have added discussion of the fact that our emotion induction had failed and that we therefore could not assess effects of emotion on DA effectiveness. We have added: “Moreover, all participants in experiment 1 felt more
sad and anxious after the induction with happy, sad or anxious emotions. The most plausible explanation therefore is that besides the three different mood induction techniques that were used in the study (a movie, music and suggestions in the script) all participants had to read a relatively sad hypothetical script and make a difficult (hypothetical) decision, which may have overruled the effect of the other mood induction techniques. Unfortunately, this precluded us from analyzing the DA effectiveness in different emotional states.

7. Please check the statistics regarding "make a decision" in table one. Is the difference between 80% & 82.8% really statistically significant?
Reviewer 2 was completely right in suggesting to check the statistics regarding “making a decision” in table 1. There was no statistical significant difference between the groups, so we have deleted the symbol for statistical significance. We have additionally checked all other analyses for possible inconsistencies. We would like to thank the reviewer for noticing this.

8. Was decisional conflict measured pre-intervention?
Decisional conflict as an outcome measure is useful only when people are unsure what they should do. When comparing groups it is therefore important to confirm that there are no differences in pre-intervention decisional conflict. Was this done? If not, this possibility should be discussed as a study limitation and the results interpreted accordingly.
Reviewer 2 is totally correct by pointing out that it would be interesting to know whether there are pre-treatment differences on important outcome measures. However, with respect to decisional conflict, we do not think that the lack of a baseline decisional conflict measure is a study limitation for two reasons. First, we think that, especially in case of a hypothetical decision, it is not interesting to measure decisional conflict when no decision has been made yet, and somebody is not yet aware of the possible options and their consequences. Therefore it is more interesting to measure decisional conflict post-decision making. Second, in both our experiments respondents were randomized, thus enabling us to assume that there are no baseline differences between the groups. These considerations made us decide not to mention the lack of a pre-interventional measure of decisional conflict as a limitation to our study. We hope that the reviewer can agree with us.

9. Information should be provided regarding the procedure used to allocate subjects to study groups.
As requested by the reviewer we have added information about the randomization of participants to the different randomization groups. We have added to the results section of experiment 1: “For the randomization we used a block randomization scheme with variable blocks sizes containing all 6 possible combinations of emotions and type of decision aid randomization per block, developed by the department of medical statistics.
of the LUMC.”. and experiment 2: For the randomization we used a block randomization scheme with variable blocks sizes containing all 3 possible combinations of type of decision aid and referral to the VCE per block, developed by the department of medical statistics of the LUMC.”.