Author's response to reviews

Title: Impact of a web-based treatment decision aid for early-stage prostate cancer on shared decision making and health outcomes: Study protocol for a cluster randomised controlled trial

Authors:

Maarten Cuypers (m.cuypers@uvt.nl)
Romy RED Lamers (r.lamers@elisabeth.nl)
Paul JM Kil (p.kil@elisabeth.nl)
Lonneke V van de Poll-Franse (L.vdpoll@uvt.nl)
Marieke de Vries (mariekedevries@uvt.nl)

Version: 2
Date: 23 April 2015

Author's response to reviews: see over
Dear Prof. Altman, Prof. Furberg and Prof. Grimshaw,

Thank you very much for the opportunity to submit a revision of our manuscript, *Impact of a web-based treatment decision aid for early-stage prostate cancer on shared decision making and health outcomes: Study protocol for a cluster randomised controlled trial*. We found the reviews to be enormously helpful. In addition to responding in detail to the reviews below, we would like to point out some of the large changes we have made to improve this manuscript.

First, as Reviewers indicated our article is mainly relevant for those in close relation to the field. Our introduction should therefore have a more specific focus on decision making in prostate cancer care, rather than the broad overview from general medicine we initially provided. We have adjusted our introduction section accordingly and included a discussion (page 5, paragraph 1) on two review articles on prostate cancer treatment decision aids (Lin, 2009; Violette, 2015). This also further supports the relevance of our study as these reviews point at the need for implementation research and the promotion of shared decision making in decision aid research.

Secondly, Reviewer 2 and Reviewer 3 addressed some concerns about the development of our decision aid. Reviewer 2 pointed out the need for a more precise description of our values clarification methods (VCMs). We thank the Reviewer for pointing this out. Instead of the two examples we provided in our initial version we now explain all our VCMs as we feel this is an important aspect of our decision aid and requires a careful explanation. Reviewer 3 discussed the adaptation process from the original Canadian decision aid to the current Dutch version. We agree with the Reviewer that we should have explained this process in more detail. Particularly, we based the structure of our DA on the Canadian original, but textual content was newly developed based on Dutch and European clinical guidelines. Urologists and communication experts reviewed the content. We have added this explanation to the introduction section (page 5, paragraph 2).

Third, Reviewer 1 and 3 asked some questions about the structure of our outcome measures. We agreed with Reviewer 3 that it would be best to identify one primary outcome, rather than having three. We have revised the organization of our outcomes measures and identified decisional conflict as primary outcome measure. From the three primary measures we previously indicated (also treatment satisfaction and decisional regret), decisional conflict is
most directly related to the decision making process and DA usage. Further, we have re-organized our secondary outcomes into either a Shared decision making outcome or Health outcome to provide a better understanding on how the different variables fit within our conceptual framework. We separately state the potentially moderating variables we are taking into account (e.g. health skills, personality) as Other measures (Table 1 and table 2). Related to the identification of decisional conflict as primary measure, we would like to address Reviewer 1’s concern about using decisional conflict, satisfaction and regret as outcomes. We thank the Reviewer for addressing this point and for referring to a relevant discussion on this topic. Although, we agree with the potential limitation of especially the decisional conflict scale and have included a section in our discussion acknowledging this potential limitation, we believe decisional conflict still represents the best available affective-cognitive outcome measure that captures the uncertainty involved to prostate cancer treatment decision making. Uncertainty about disease progression, treatment effectiveness and side-effect impact are key elements of the decision making process in prostate cancer care. Preliminary investigations prior to the current study taught us that decisional conflict levels are substantial and is a major cause of distress. We expect that our DA is able to reduce this experienced distress. For meaningful interpretation of our effects we also have additional outcome measures included in our study, next to decisional conflict, that can further support our findings or can indicate bias if present. Further, to our knowledge no other measure is available to capture this specific feeling of distress during treatment decision making.

Fourth, there were a number of Reviewer comments that caused us to believe that is was necessary to restructure our statistics section. We understand the confusion the first version could have caused. We adjusted this section according to the Reviewers’ suggestions (page 15). Additionally, Reviewer 3 noticed that our references for the estimations of our effect size and ICC were not specific for prostate cancer studies or cluster RCTs. We thank the Reviewer for pointing this out. We have reviewed the available literature again and have added more precise sources (page 14, paragraph 2). These additional sources further supported our initial approach to make conservative estimations as a large variation of effect sizes and ICCs are reported.

We believe that these changes, in addition to the changes described below, address the comments and concerns of all the Reviewers. We would like to thank you and the Reviewers for the very helpful comments and suggestions based on the previous version of our manuscript. We believe that based on these comments, the quality of the manuscript has substantially improved. We hope that the manuscript now meets the high standards for publication of Trials, and look forward to learning of your decision in relation to our manuscript.

Sincerely Yours,
also on behalf of Romy Lamers, Paul Kil, Lonneke van de Poll and Marieke de Vries,

Maarten Cuypers
Author reply to Reviewer comments

Comments from the editor

1. Please ensure the title conforms to journal style for study protocol articles. The title should follow the format ___________: study protocol for a randomized controlled trial.

   We have adjusted the title accordingly into:
   **Impact of a web-based treatment decision aid for early-stage prostate cancer on shared decision making and health outcomes: Study protocol for a cluster randomised controlled trial**

2. Please include the full names and email addresses of all authors on the title page.

   *These details have now been added.*

3. Please include the date of registration with the trial registration number at the end of the Abstract.

   *This information has been added:*
   The Netherlands National Trial Register NTR4554, **registration date May 1st, 2014**.

Reviewer #1

1. I am fully aware that decisional conflict, treatment satisfaction and decisional regret are pretty standard outcomes measures, but they actually do an extremely poor job of distinguishing good from bad decisions. See, for example, Fagerlin Medical Decision Making [http://www.ncbi.nlm.nih.gov/pubmed/17873251](http://www.ncbi.nlm.nih.gov/pubmed/17873251). In brief, it is not at all clear that lower conflict indicates better decision-making (e.g. a very biased decision aid would lead to lower conflict, because patients would be more sure that they had made the right decision). And regret is extremely time sensitive (e.g. a man may have no regrets 12 months after choosing active surveillance compared to a surgery patient enduring incontinence after surgery; 10 years later, when the active surveillance patient is dying from metastatic disease, things may be very different). At the very least, the authors need to acknowledge the very important limitations of these endpoints and to discuss how interpretations of the study may be accordingly difficult.

   *We thank the Reviewer for addressing this concern and referring to relevant additional literature on this matter. We agree with the Reviewer that we should have addressed this concern. As the Reviewer also recognizes, decisional conflict, satisfaction and regret are common measures in decision aid*
evaluations. For decisional conflict, which is our primary outcome, we now explicitly refer to other studies using the DSC as an indication of its wide application and acceptance:

*(Page 11)* The DSC is widely accepted and applied as main outcome measure in (PrCa) DA trials [34] [24] [35] [36].

Further, we address this matter in our discussion section where we acknowledge the limitations Reviewer 1 referred to:

*(Page 18)* Our trial has defined decisional conflict as primary measure. As previously mentioned, the DCS is a widely accepted and applied measure in DA evaluations. However, the DCS is also subject to some discussion in the literature about its usefulness as outcome measure in DA evaluations [68]. This is mainly due to its limitation to identify a good decision as a person’s underlying sensitivity to uncertainty may not be fully represented in a high or low decisional conflict score. For example, a high score on the DCS could also represent the effort that one takes to be involved in the decision making process and absorbing all available information and therefore becoming aware of the difficulty of the decision. Although we are aware of this potential limitation of the DCS, we believe decisional conflict represents the best available affective-cognitive outcome measure that captures the uncertainty involved to prostate cancer treatment decision making. Uncertainty about disease progression, treatment success and side-effect impact are key elements of the decision making process in prostate cancer care. Preliminary investigations prior to the current study taught us that decisional conflict levels are substantial and we expect that our DA is able to reduce these levels and that this potential reduction is meaningful. For meaningful interpretation of our effects we also have additional outcome measures available which can support our findings or can indicate bias if present. Many of our secondary measures focus on the decision making process (knowledge, satisfaction with information provision, decision making role) rather than the outcome in terms of a ‘good’ or ‘bad’ decision, this ensures that our conclusions on the usefulness of the DA will not solely depend on interpretation of the DCS.

While we also fully agree with Reviewer 1 on the time and treatment sensitivity aspects in decisional regret and treatment satisfaction, we believe that the design of our study allows for taking this concern into account. Between our intervention and control group we can compare regret and satisfaction levels based on selected treatment and thus control for bias based on treatment choice. The time aspect is captured by offering questionnaires at a fixed moment in time for all treatment groups. This means that regret and satisfaction levels may variate between treatment groups at the same measurement, however within each treatment group every participant is at the same stage since treatment started. Additionally, we check if any follow up treatments were started since the initial treatment choice.

2. The statistics section is confused and vague. This is a cluster randomized trial, and so analysis must take account of the design. However, the authors have a long section describing a number of different possible analyses (e.g. t-tests or non-parametric equivalents) that do not reflect the cluster design. They then briefly describe a multilevel modeling approach without giving many details at all. The general principle of a statistical section in a protocol is that two statisticians working independently would write a similar results section on the basis of the
protocol alone. The statistics section should describe each hypothesis tested and then give a clear and unambiguous description of the statistical methods.

We understand the confusion caused by the previous version of the statistics section and would like to thank the Reviewer for highlighting this. We made changes according to the Reviewers’ suggestions:

(Page 15) All analyses will be conducted using SPSS version 19.0 (Statistical Package for Social Sciences, Chicago, IL, USA). A 0.05-significance level will be adopted in all statistical tests.

We will perform a descriptive statistical analysis of organizational (hospitals) and socio-demographic (patients) characteristics in order to assure the comparability of the intervention and control groups. Baseline measures and changes in outcome variables over the study period for each study arm will be presented as means (± SD).

The main outcome decisional conflict is measured at T1 and will be compared between both groups (intervention and control). Multilevel modelling will be carried to take the hierarchical structure of the data into account by specifying random effects at both hospital and patient level. The least square mean proportions will be estimated and compared to assess the effect of the DA on decisional conflict.

The secondary outcomes will also be compared between both groups using multilevel modelling. Some of the secondary measures consist of repeated measures (e.g. HRQoL, decisional regret) and will be treated according the appropriate mixed-model approach, that is repeated measures anova/ancova for outcomes with two time points (decisional regret, treatment satisfaction) and a random coefficient approach for outcomes with three time points (HRQoL) [67]. Observed variation in treatment choice during the trial period will be compared between groups and at individual hospital level. For this second comparison each hospital’s particular historical treatment variation profile (2008-2012) is obtained from the Netherlands Cancer Registry.

Potentially confounding variables (e.g. personality, health skills, age) will be explored for their impact on the primary and secondary outcomes. Missing data and drop-out will be described.

3. The EORTC30 includes questions such as “Do you have any trouble doing strenuous activities, like carrying a heavy shopping bag or a suitcase?”; “Have you vomited?” and “[Have you felt] irritable?”. On what possible grounds would we expect patients to answer such questions differently depending on their group assignment?

We agree with the Reviewer that some of the items within the EORTC30 do not appear to be relevant when evaluating a decision aid. Especially for specific questions from the EORTC30 like Reviewer 1 highlights, we do not expect a difference based on assignment to either intervention or control group. However, we do expect a consequence for HRQoL in general if the treatment decision follows DA usage, expecting that this better supports SDM and the treatment decision is then more in line with personal preferences. To capture HRQoL different scales are available, of which some are less extensive than the EORTC30. To be consistent with previous measures by our group and to be able to include the potentially more relevant prostate specific PR25 module (for which the manual dictates that it is accompanied by the EORTC30) we decided to include the EORTC30 as well. Also, reviews on SDM and DAs indicate a need
for HRQoL to be investigated more as outcome of SDM and DA interventions to link empirical measures of SDM and patient behavior to health outcomes (e.g. Lin, 2009; Shay, 2014).

Reviewer #2

This is a well-designed study with appropriate outcomes and a sound statistical design.

We would like to thank Reviewer 2 for this compliment on our work.

Major revisions: The authors note that decision aids have struggled to find a place in clinical practice due to problems with integration into clinical workflow. At the same time, details of how they propose to integrate this DA into clinical practice are scant in the methods, and apparently they plan to leave integration up to each clinic. A clear template of at least one way to identify patients prior to their cancer consultation, details on who is expected to identify these patients and send them a card with relevant clinical data, and ways to overcome implementation barriers might be helpful.

We understand that our initial implementation procedure could be interpreted as if there was no guidance for introducing the DA to patients. Our intention was to reduce the extra workload that introducing the DA could involve. To achieve that, development of the DA included cooperation with urologists to let the DA fit with clinical practice. The time between diagnosis and (standard) information provision was identified as the ideal moment to introduce the DA to the patient. As this exact process differs per hospital, the pragmatic aspect of our trial allows hospitals to implement the DA at the most suitable moment. However, in daily practice this means that patients either receive the DA at diagnosis (from their urologist) or in the follow-up consultation with a (oncology) nurse. Since these are the only two routines that occur, we now describe that the DA is offered at diagnosis by the urologist or the (oncology) nurse if there is a follow up meeting between diagnosis and decision making:

(Page 8) After diagnosed with prostate cancer, but before a treatment decision is made, patients in the intervention arm receive access to the online DA. Healthcare providers are instructed to introduce the DA and the study at diagnosis. However, the pragmatic nature of this trial allows hospitals to integrate the introduction of the DA with their standard information provision routines if that follows later due to follow-up diagnostics or an additional consultation with (oncology) nurses. In daily practice this means that it is either the urologist or the (oncology) nurse who introduces the DA to the patient. To use the DA, patients receive a card with their relevant disease characteristics (PSA, Gleason, and eligible treatment options) and a personal username and password to gain online access to the DA. If a nurse introduces the DA and accompanying access card the urologist should provide the requested clinical characteristics to the nurse, either by filling in the card or by leaving a note in the patients’ record.
Discretionary Revisions:

Very little is said about the development of the DA. I assume it was done with the input of patients, but the values clarification exercise is vaguely described. How is a patient supposed to understand the choice between: "I find it important that all cancer cells are removed from my body versus I find it important that the cancer cells die and not grow further"? It sounds like 'cure' is the patient centered outcome being evaluated here; were these 'paths to cure' identified as important to understand in Dutch patients? Are they indicating agreement via Likert scale or some other method? Are there ways to compare across different domains of treatment in these exercises (say, risk of urinary control vs risk of cure)?

The values clarification methods VCMs are an important aspect of our DA, we therefore thank the Reviewer for indicating that our initial explanation was not adequate. Instead of the two examples we included in the original manuscript, we now added all VCMs and the topic or domain they cover. We also explain the response method in more detail:

(page 9) The DA offers a stepwise guidance through the decision process. In the first step, general information about prostate cancer is provided. The second step offers the consideration between active surveillance and curative treatment options (surgery or radiotherapy). Values clarification statements are presented in this step to elicit a patient’s preference toward active surveillance or curative treatment based on three main differences between AS and curative treatment; acceptance of deferring treatment (‘I am confident enough that I will be treated on time, if necessary’ versus ‘I do not want to postpone treatment because I do not want to be too late’), avoiding possibly unnecessary treatment (‘If treatment might be unnecessary, I would rather wait’ versus ‘I prefer treatment, even if it might be unnecessary’) and the acceptance of treatment side effects (‘I find possible treatment side effects like erectile and urinary dysfunctions difficult to accept’ versus ‘I find the possible treatment side effects acceptable’). Each statement is related to one of the two offered treatment alternatives in this step. On a slider scale, patients can indicate for each set of statements, the strength of their preference towards one of the alternatives.

Following the same structure, the next step supports the consideration between surgery and radiotherapy. For surgery, three common methods are discussed (laparoscopic, open and robot assisted). For radiotherapy this consists of brachytherapy and EBRT. Again, information provision is followed by values clarification statements. The VCMs in this step emphasize the main differences between surgery and radiation therapy (both brachy and EBRT) in terms of treatment procedure (‘I find it important that all cancer cells are removed from my body versus I find it important that the cancer cells die and not grow further’), side effects (‘I find bowel problems worse than incontinence’ versus ‘I find incontinence worse than bowel problems’), secondary treatment (‘I am comforted by the thought that I can have radiation if surgery is unsuccessful’ versus ‘I accept that surgery is difficult after radiation’) and fear for surgery (‘I am not anxious for surgery’ versus ‘I am anxious for surgery’). If a patient already indicated a preference for active surveillance in the previous step, the program allows patients to ignore this part and continue to the last step. As a conclusion patients are asked to indicate their final treatment preference and briefly explain their choice. The DA does not provide a treatment advice, but helps the patient to reach a decision. A summary then provides an overview of all
answers to the statements and the patients’ final preference. To discuss this summary with their urologist, the summary can be printed or accessed online during the next consultation.

All statements used in the VCMs were developed by a team of urologists, psychologists and engineers based on previous experience and observation of conversations where treatment decisions were discussed. The statements were evaluated during usability-testing among patients, urologists and nurses (N=10).

Reviewer #3

1. This protocol paper is focused on an important and timely topic. However, many issues detract from its quality. Many of the references in the background section are "one-off" citing studies (orthopedics, internal medicine) instead of a careful review of more pertinent trials in decision making for prostate cancer.

We agree with the Reviewer that the study will mainly be of interest to a specific audience. Therefore, we decided not to give a broad and general introduction based on the wide application of SDM and DAs in different fields of medicine. We have removed some of the general introduction and focused more on prostate cancer treatment decision making by discussing two review articles on prostate cancer decision aids that have included many of the relevant PrCa DA trials of the past years (Lin 2009, Violette, 2015).

(Page 5) During the past decade, several decision aids (DAs) have been developed with special focus on prostate cancer care. Instruments range from information booklets to tailored web-based tools. The variety in used formats may have contributed to the finding that effects on decisional outcomes have been inconsistent across randomized trials and that no effects on choice have been found [21] [22]. Systematic reviews further emphasize that many previous studies are at high risk of selection bias due to inadequate concealment or blinding of data collectors and outcomes assessors, and that more studies are needed to determine how DAs can be implemented best in clinical practice [21] [22].

Determining effect of a DA intervention and finding optimal implementation methods are both aims of the current trial. A web-based prostate cancer treatment DA was developed to fit with Dutch clinical workflow. Based on the structure of an existing DA developed by Feldman-Stewart and colleagues [23] [24], Dutch content was written and values clarification methods (VCMs) were added. Adaptation of the DA was based on the International Patient Decision Aid Standards (IPDAS) [25].

2. The translation from Canadian English to Dutch was given cursory attention. Simple direct translation does not assure a culturally sensitive and linguistically appropriate product.

We understand the initial explanation on the development of the DA could have caused some confusion about translation practices. Although we based our DA on an existing Canadian DA, the translation process only addresses the topics included in the DA. The actual content was newly written, based on European and Dutch clinical guidelines. A prior full translation did indeed reveal the same
issues that Reviewer 3 highlights and led to the development of new Dutch content. We have now explained the development more thoroughly, in addition to previous Reviewer’s comment on the development of the VCM’s.

(page 5). A web-based prostate cancer treatment DA was developed to fit with Dutch clinical workflow. Based on the structure of an existing DA developed by Feldman-Stewart and colleagues [23] [24], new Dutch content was developed and values clarification methods (VCMs) were added. The final version of the DA was examined in a usability test that consisted of patients, urologists, nurses and a radiotherapist (N=10). All textual content was reviewed by urologists (medical accuracy) and communication experts (readability). Adaptation of the DA was based on the International Patient Decision Aid Standards (IPDAS) [25].

3. The outcome measures are extensive and certainly may be burdensome. What is the conceptual framework and how do the many variables fit?

We are aware of the length and impact of the questionnaire to our responders. However, our group has experience with similar studies with questionnaires of comparable length. The number of variables fit with the scope of this trial that investigates both effect and implementation of the DA. This means we need measures to capture direct decision outcomes (treatment choice), evaluation of the decision making process (decisional conflict, decision making role), DA-evaluation (knowledge, satisfaction with information provision) and long-term outcomes that could be affected by a treatment choice that is more in line with personal preferences and characteristics (regret, HRQoL). For optimal implementation we investigate factors that might stimulate or hinder DA-usage, therefore we added the personality variables (e.g. NFC, optimism), capability measures (numeracy, literacy) and subjective health evaluations (health status and side-effect impact). We expect to identify subgroups within our population that benefit more from the DA than others.

4. Will the study design adequately test the hypothesis?
   • Doubtful. The sample size assumes an effect size of .5 with no references to results of previous randomized trials of decision aids for localized prostate cancer (e.g., Feldman-Stewart, 2012 or Berry, 2013).

We thank the Reviewer for pointing this out, and for referring to useful comparable trials. Although previous trials with comparable outcomes have been published, all have their specific intervention and population. A recent review reported effect sizes for decisional conflict in PrCa DA trials ranging from 0 to .6 (Violette, 2015). Given this variance in effect sizes and the characteristics of our trial (novel web-based DA, cluster RCT, Dutch population) no appropriate comparison is available. In the absence of an appropriate comparison we calculated our sample size with the assumption to at least identify a clinically relevant effect size (.5). If powered at a smaller effect size it is doubtful if significant differences then remain clinically relevant.
5. Three primary measures are identified; which is the real primary outcome?

From our initial three primary outcomes (decisional conflict, treatment satisfaction, decisional regret), decisional conflict is most directly related to the decision making process and DA usage. Therefore we defined decisional conflict as our primary outcome.

6. Secondary outcomes are not well organized.

We hope to have clarified this by restructuring the paragraph that discusses the outcome measures:

(Page 12) **Secondary outcomes**

Secondary outcome measures can be categorized as either shared decision making outcomes or health outcomes. Shared decision making outcomes consist of decisional regret [37], perceived and preferred decision making role [38], and preparation for decision making to assess a patient’s preparation for decision making and dialoguing with his clinician [39]. Furthermore, a single-item question will evaluate the perceived patient-doctor relationship and the development of this relationship over time. Also, satisfaction with information provision [40] and knowledge [41] will be assessed.

Health outcomes refer to the actual treatment choice and any changes in treatment preference during the decision making process. Treatment satisfaction will be measured with a single-item question: ‘Are you satisfied with the way your treatment was or is executed?’ To assess health-related quality of life (HRQoL) the EORTC QLQ-C30 [42] will be used. This questionnaire is developed specific to assess HRQoL in cancer patients. Much of the content of the questionnaire is appropriate for extended monitoring of health status, including scales assessing physical, role, cognitive and emotional functioning, fatigue and sleep problems, and overall health and quality of life. This core instrument is supplemented with the prostate cancer-specific HRQoL questionnaire EORTC QLQ-PR25 [43]. This 25-item questionnaire assesses urinary, bowel and sexual symptoms and functioning, and the side-effects of hormonal treatment, though hormonal treatment is not offered as initial treatment in this study’s sample. Health outcomes are further assessed by means of an evaluation of side-effect impact [44] [45] and health status acceptance and subjective control [46].

<table>
<thead>
<tr>
<th>Table 1 Outcome measures</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Shared decision-making outcomes</strong></td>
</tr>
<tr>
<td>Decisonal conflict</td>
</tr>
<tr>
<td>Decisional regret</td>
</tr>
<tr>
<td>PrCa Knowledge</td>
</tr>
<tr>
<td>Satisfaction with information</td>
</tr>
<tr>
<td>Decision making preparedness</td>
</tr>
<tr>
<td>Decision making role</td>
</tr>
<tr>
<td>Perceived doctor-patient relationship</td>
</tr>
</tbody>
</table>
Actual treatment choice and health outcomes

<table>
<thead>
<tr>
<th>Measures</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial preference and treatment choice</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Treatment satisfaction</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health-related Quality of Life</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Side-effect impact</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Acceptance &amp; Control over health status</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Table 2 Other measures

<table>
<thead>
<tr>
<th>Measures</th>
<th>Instrument</th>
<th>T1</th>
<th>T2</th>
<th>T3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Implementation (intervention only)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DA-Acceptability</td>
<td>Self-developed</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Health skills</td>
<td>Decision self-efficacy scale</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>HRS Experimental numeracy module</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Health numeracy</td>
<td>STOHFLA-brief</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychosocial variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety and depression</td>
<td>HADS, PC-max</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Personality</td>
<td>LOT-R, BFI-10</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Information seeking preferences</td>
<td>API, NFC-short, Maximization scale</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Sociodemographics and other healthcare utilization</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

7. The cluster randomized trial is not adequately justified with institution as the unit of analysis. What about differences between clinicians? If Dutch clinicians have no variability in SDM, perhaps this could work. Doubtful that the study "training would change years of practice behaviors.

We thank the Reviewer for highlighting this inaccuracy. We fully agree with Reviewer 3 that variability within hospitals is also likely to exist. In our discussion section we added our considerations on this topic and explain why we have not included individual clinicians as unit of analysis:

(Page 17) Although we are aware of the fact that individual differences between clinicians could also affect decision outcomes, there are some considerations that justify taking institution as the unit of analysis. First, diagnosis and offered treatment plans are often the result of multi-disciplinary consideration (e.g. urologists, radiotherapists, oncologists). Secondly, specialization often leads to some clinicians seeing the majority of PrCa patients within an institution. Taking the clinician as unit of analysis could lead to too small clusters in some cases. On the other hand, clinician specialization could also lead to patients visiting multiple clinicians within a hospital before a final decision is made, making it difficult to attribute a treatment decision to a certain clinician. Third, information provision and decisional support is often provided by specialized (oncology) nurses. Typically they assist
more than one clinician which could contaminate individual clinicians’ data. Finally, regional variation in treatment practices, which is expected to be influenced by DAs as explained in previous sections, is generally reported at hospital level. This indicates that there are influences at hospital level driving practice variation that go beyond individual differences between clinicians within a hospital. The reported variation in selected treatments between hospitals is available for hospitals included in our study, though no data is available on individual clinician’s variability.

8. The ICC of .1 is based on studies in general medicine, neither of which were cluster randomized trials.

We thank the Reviewer for this comment. First we should correct that the ICC used in sample size calculation was not .1 but .01. In addition to Reviewer 3’s previous comment about our estimation for the effect size, also no appropriate ICC estimation could be obtained from previous trials. However, a more relevant reference was found, which also advises a conservative estimation due to large variations in reported ICCs among different studies with the same outcome measures. Our selected ICC and other assumptions used for our sample size calculation should therefore also be considered as conservative. In the sample size section we included this explanation:

(Page 14) ICC ranges from 0 to .1 are considered common in medical literature [62]. A more specific review of ICC values in (cluster) RCTs with psychosocial measures is provided by Bell and McKenzie [63], which also included a cluster RCT evaluating a group support tool for prostate cancer patients [64]. The median estimated value for 82 longitudinal ICCs from 15 included studies was 0.0007, and the range found for decisional conflict was between 0 and 0.02. Given the considerable variability in ICCs that is found in literature, ICC for the current trial is set conservative at 0.01.

9. Program access requires Internet connectivity, limiting the sample ad of course then, the findings. What devices are allowed?

We thank the Reviewer for highlighting this point. It is important to mention that our DA is compatible with all devices (PC, laptop, tablet, smartphone). Internet access in The Netherlands is among the highest in the world, 94% of the households in The Netherlands has internet access (2013, Worldbank). Even in older age groups (65-75 years) regular internet usage is at 80% (2013, Statistics Netherlands). Internet is used to find additional information and healthcare providers refer to webpages as part of standard information provision. Therefore we do not expect that this leads to biased sampling or findings. Additionally, since we mainly distribute questionnaires digitally, the control group also consists of men familiar with internet usage. Since all devices with internet connection are allowed, we explicitly added this to the inclusion criteria:

- (Page 8) Patient has access to a PC, laptop or tablet with internet connection

We have also added a section to our discussion about this point:

(Page 18) A potential limitation of our DA is that a device with internet connection is needed to use the DA, which could affect our sample and consequently our findings. Although
we are aware that this could be a relevant issue in many regions in the world, we do not expect biased results in our trial. The Worldbank has estimated internet access in The Netherlands is among the highest in the world, with 94% of the households (2013) having internet access (www.worldbank.org). Even in older age groups (65-75 years) regular internet access is at 80%, and this percentage is rapidly increasing (2013, Statistics Netherlands). Internet is routinely referred to as part of information provision in standard care. As most of our questionnaires (in both groups) are sent via email, internet access and the ability to use it is also required in both groups, assuring group comparability on this matter.

10. Are sufficient details provided to allow replication of the work or comparison with related analyses: if not, what is missing?
   • Probably

   After careful review we indeed believe to have provided sufficient details for replication of our work.

11. Is the planned statistical analysis appropriate?
   • The analysis section is confusing and does address all aims or measures.

   We would like to thank the Reviewer for pointing this out. As outlined above (Reviewer 1, point 2) we substantially revised the section on planned statistical analyses.