

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form ([see an example](#)) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	The INvolvement of breast CANcer patients during oncological consultations. A multi-centre randomized controlled trial. The INCA study protocol.
AUTHORS	Goss, Claudia; Ghilardi, Alberto; Deledda, giuseppe; Buizza, Chiara; Bottacini, Alessandro; Del Piccolo, Lidia; Rimondini, Michela; Chiodera, Federica; Mazzi, Maria Angela; Ballarin, Mario; Bighelli, Irene; Strepparava, Maria Grazia; Molino, Annamaria; Fiorio, Elena; Nortilli, Rolando; Caliolo, Chiara; Zuliani, Serena; Auriemma, Alessandra; Maspero, Federica; Simoncini, Edda; Ragni, Fulvio; Brown, Richard; Zimmermann, Christa

VERSION 1 - REVIEW

REVIEWER	Mary Ann O'Brien, PhD Assistant Professor, Department of Family and Community Medicine, University of Toronto, Canada I do not have any competing interests to declare.
REVIEW RETURNED	18-Dec-2012

THE STUDY	<p>Adequacy of the overall research design Review Summary: A randomized controlled design is appropriate to answer the research question. However, the research hypotheses (pages 6 and 7) are not clear. With respect to the intervention, it is uncertain if patients and any accompanying persons will have adequate time to review the intervention, a question prompt sheet (QPS). Several of the outcome measures do not appear to be directly related to the primary study aims or hypotheses. For example, "consultation atmosphere" (page 15), measured with both the Roter Interaction Analysis System and the Assessment of Interpersonal Motivation in Transcripts is not well-linked to a specific study hypothesis. From a design and analysis perspective, the authors do not appear to take into consideration the potential effect of clustering in their primary analysis even though it appears that patients are nested within oncologists who are nested within centre. Additional details are found in the sections that follow.</p> <p>Study Participants Abstract page 2; Figure 1, page 8; Sample and recruitment, page 11. The protocol does not specify the gender of the patients in the Abstract, in Figure 1, nor in section on Sample and recruitment. In several places, the investigators refer to 'her' or 'his/her' making it difficult to ascertain the gender. Study information found at clinicaltrials.gov indicates that the patients are female; this information should be included in both the abstract and in the section on the sample.</p> <p>Description of Methods</p>
------------------	--

Recruitment. There is insufficient description of the recruitment process.

Adequacy of the Intervention. It is unclear how the QPS intervention is given to the patients and how much time the participants have to review 50 questions. In addition, the QPS is collected by a project staff member prior to the consultation. Presumably this was done so that oncologists would be blind to the intervention or control status of the patient. However, the QPS as originally designed was used by the patient within the consultation to help them ask relevant questions (Butow et al. 1994). The authors should explain how patients will be able to remember all the questions they wanted to ask when the prompt sheet is taken away prior to the consultation.

Outcome measures. As stated in the review summary, the authors should better justify the inclusion of outcome measures that are not directly related to the study aims/hypotheses.

With respect to self-reported process and outcome questionnaires, it is possible that patients and any accompanying persons may experience considerable burden in completing so many questionnaires some of which appear to be tangential to the study (e.g., the 24-item subscales of Eysenck Personality Questionnaire).

Statistical Methods. There are several concerns about the statistical methods. There is insufficient detail to judge the adequacy of the planned analyses. The authors state that they will use t tests in the primary analysis. Do the authors believe that the data will be normally distributed with equal group variances? In one of the references cited by the authors of previous research on the QPS, the data were not normally distributed (Brown et al.). Moreover and as stated in the review summary, the authors do not appear to take into consideration the potential effect of clustering in their primary analysis (Donner and Klar, 2000) even though it appears that patients are nested within oncologists who are nested within centre. In their secondary analysis, they are proposing multi-level modelling which may consider clustering. It is unclear why t-tests are used in the primary analysis and multi-level modeling in the secondary analysis. In addition, the authors state on page 18 lines 27-30 that the primary outcome, "significant increase of patient questions, will be compared in the two arms". It is not clear if the authors mean that the primary outcome is a pre-post difference score that will be compared between intervention and control groups or if the primary outcome is the post intervention difference between intervention and controls groups in the total number of questions.

Standard of written English. On occasion, the authors use imprecise language which does not convey the presumed meaning. One example is the use of the word "arguments" rather than "questions" in the instructions to the patients. In many instances, there are typographical errors which are distracting for the reader. For example, see the complete list of authors and Reference 1 in the reference list. There are other issues that are related to writing style. For example, Table 1 does not contain explanations of the short forms so the reader needs to flip back and forth between the Table and the text to find the names of the outcome measures. Table 1 could be improved by including the full name of the outcome measure along with the key constructs measured.

Article Summary

There is mismatch between the abstract and the summary with

	<p>respect to the outcome of patient satisfaction and coping. Study limitations are not adequately discussed. The last bullet point in the article summary should be deleted. The QPS as tested in this study is not brought into the consultation.</p> <p>Abstract Introduction: The authors need to be specific about which aspect of patient involvement they believe is important e.g., in decision making?, in the consultation? The first two sentences are not linked. The second study aim relates to the family member but there are no outcomes listed in the abstract that pertain to this person.</p> <p>Methods and Analysis: The primary outcome should be stated. As above, none of the outcomes or the hypotheses is related to the family member. Another relatively minor issue is that the authors use the term ‘family member’ as well as the terms “accompanying person” or “key person”. It would be helpful if the authors could use one term throughout.</p> <p>References Two references are missing. The details are included below. van der Meulen N, Jansen J, van Dulmen S, Bensing J, van Weert J. Interventions to improve recall of medical information in cancer patients: a systematic review of the literature. <i>Psychooncology</i>. 2008 Sep;17(9):857-68. van Weert JC, Jansen J, Spreeuwenberg PM, van Dulmen S, Bensing JM. Effects of communication skills training and a Question Prompt Sheet to improve communication with older cancer patients: a randomized controlled trial. <i>Crit Rev Oncol Hematol</i>. 2011 Oct;80(1):145-59. doi: 10.1016/j.critrevonc.2010.10.010. Epub 2010 Nov 13.</p>
<p>GENERAL COMMENTS</p>	<p>Additional Comments Page 6 Lines 54 and 57 and Page7 Lines 3-25. It is quite difficult to assess the logic of this section as it is written. For example, it is not clear what the authors mean by “a different perception by the oncologist and by the patient of the doctor-patient relationship“. For each hypothesis it would be clearer if the authors included the relevant outcome measure and the comparator. Furthermore, the authors should include more details of how they plan to measure if the “question prompt sheet intervention (extended to the accompanying key person) changes the key person’s role and participation during the consultation”.</p> <p>Minor Comments Page 7 Line 52. The authors state that the oncologists complete two measures post consultation but only one is included in Table 1. Page 9 Table 1. CPS. Indicate the number of vignettes and how each is scored. It is unclear when the DP is scored by oncologists. The audiorecording is not a scale but a data collection tool. The OPTION scale does not measure involvement level. It measures health professional behaviours intended to involve patients (Elwyn et al. 2005). The footnote indicates that certain questionnaires were adapted for companions. The authors need to provide more detail about how this adaptation was done. Page 11 Line 43. The authors state that the patients and companions complete six questionnaires preconsultation but Figure 1 indicates three. Page 11 Line 23. The authors should explain how early stage breast cancer is defined.</p>

	<p>Page 11 Lines 56-59. The authors state that patients and companions complete six questionnaires post consultation but Figure 1 indicates five questionnaires. Further they state that projects staff will assist in questionnaire completion. What does 'assist' mean and how will bias be avoided?</p> <p>Page 12 Lines7-8. The authors need to justify including a questionnaire in which oncologists rate the "difficulty" of the patient and indicate its relevance to the study aims or hypotheses. Further, the authors state on Page 15 that they added an additional four items. What are the psychometric properties of this revised questionnaire?</p> <p>Page 12 Lines 27,28. Why have the authors indicated the use of Assessment of Interpersonal Motivation in Transcripts in guiding "non verbal behaviours" during the interaction when the authors are using audiotapes?</p> <p>There is considerable overlap between the text and in Table 1. As suggested previously, an expanded Table 1 would be helpful and would reduce the text description.</p> <p>Thank you for asking me to review this manuscript.</p> <p>References used in this review</p> <p>Butow PN, Dunn SM, Tattersall MH, Jones QJ. Patient participation in the cancer consultation: evaluation of a question prompt sheet. <i>Ann Oncol.</i> 1994;5:199-204.</p> <p>Brown RF, Butow PN, Dunn SM, et al. Promoting patient participation and shortening cancer consultations: a randomised trial. <i>Br J Cancer</i> 2001;8:1273-1279.</p> <p>Clayton JM, Butow PN, Tattersall MH, et al. Randomized controlled trial of a prompt list to help advanced cancer patients and their caregivers to ask questions about prognosis and end-of-life care. <i>J Clin Oncol</i> 2007;25:715-723.</p> <p>Donner A, Klar N. Design and analysis of cluster randomization trials in health research. New York, New York: Oxford University Press, 2000</p> <p>Elwyn G, Hutchings H, Edwards A, et al. The OPTION scale: measuring the extent that clinicians involve patients in decision-making tasks. <i>Health Expectat</i> 2005;8:34-42.</p>
--	---

REVIEWER	<p>Phyllis Butow Professor of Health Psychology, School of Psychology University of Sydney, NSW 2006 Australia.</p> <p>I do not have any competing interests in relation to this protocol.</p>
REVIEW RETURNED	19-Dec-2012

THE STUDY	<p>The protocol lacks a clear set of hypotheses, that address the multitude of assessments being collected.</p> <p>The statistical analysis section refers to the main study question (do</p>
------------------	---

	<p>questions increase with the use of a QPS), but do not address many other aspects of the study design. Further, the comment about interaction analysis involving sequence analysis is not related to the measures being used, and not explicated sufficiently. This may be an aspect of the ROTER analysis, but this needs to be clarified.</p>
<p>GENERAL COMMENTS</p>	<p>This is an interesting protocol evaluating the impact of a question prompt sheet for breast cancer patients attending their first visit with an oncologist. The authors claim that this is the first such study in Europe, but I am aware of several studies like this which are underway (but not yet published) in Holland, at the AMC.</p> <p>There are many strengths to this protocol, including the careful randomisation and blinding procedures, a clear sample size calculation and careful measurement of outcomes.</p> <p>I have a few queries about the protocol.</p> <p>As noted above, I think the protocol needs a series of clear hypotheses to relate the measures to the study questions. The analysis needs to be more detailed to address the hypotheses.</p> <p>I think the provision of a blank sheet to both control and intervention arms to write down the issues they would like to discuss in the consultation is a good way to control for extra attention, but may reduce the effect size of the intervention compared to control since it will trigger pre-planning in control subjects also. Perhaps the authors need to consider this in the effect size. Have they piloted this approach? (nb issues is a better word than arguments, and should be changed in the protocol. In English, arguments implies that the patient will make a series of points to argue a position.)</p> <p>In the protocol, it suggests that both the blank sheet and the QPS with ticked questions will be collected. Will the patient be left a copy? In previous trials, patients were able to bring the QPS into the consultation to refer to, and it may be less effective if this is not the case here.</p> <p>Regarding the measures, there was a lot of confusion in the protocol, with different measures being mentioned at different points - these need to be standardised. For example, several measures were not included in the Table of measures, but were referred to later, such as the CPS - pt and the CPS- onc.</p> <p>The main outcome is questions, but it was not clear to me whether this was only patient questions, or companion questions also. I would recommend both, but analysing both separately.</p> <p>The questions asked in the consultation will be compared against those ticked prior to the consultation, which is very interesting. However I do not think this comparison can be used to measure unmet information needs, as is suggested in the protocol. The patient or companion may not have asked those questions because the oncologist had already answered them. Unmet information needs could be assessed by the additional questions added to the recall measure (I think), or a much more careful analysis of the information provided and the questions asked would need to be undertaken.</p> <p>It was also not clear to me whether companions would be completing demographics, or only the patient. Gender was not listed</p>

	<p>as one of the variables being elicited from patients and companions, and I would think this was important. If the companion is a key part of the study, it may also be useful to measure the companion's relationship with the patient (e.g. spouse, daughter or son, friend etc) as their role would vary accordingly.</p> <p>There are many measures and it is not clear to me (in the absence of clear hypotheses) how each is intended to be used. Some of the measures I am not familiar with, and it may be important to do more to demonstrate their psychometric properties. The recall questionnaire is based on earlier studies, but does not seem to be related to the methods used in those studies, which were to assess recall by interview with careful prompts. It was also not clear to me how recall would be scored - the authors have added some additional questions related to patients' perceptions of the adequacy of information provision. Will these be scored separately? With all the measures, I do not see a measure of satisfaction with the consultation (just with the decision). Perhaps a general satisfaction measure is required? Does the study have adequate power to explore all of these secondary outcomes?</p> <p>Measures are being collected only immediately after the consultation. Therefore no long-term impacts are being measured. In previous studies anxiety was raised immediately after the consultation but quickly dissipated. Is it worth considering longer term outcomes?</p> <p>I do wonder whether all the methods of analysing the tapes will be needed - are the interpersonal motivation systems really related to question asking? Again, clear hypotheses would help here. There will be a LOT of work required to apply all these coding systems to the audio taped consultations.</p> <p>Are oncologists happy to complete forms after every patients is seen? I think there may be a lot of missing data but perhaps the research team is not finding this.</p>
--	---

VERSION 1 – AUTHOR RESPONSE

Reviewer: Mary Ann O'Brien, PhD

Assistant Professor, Department of Family and Community Medicine, University of Toronto, Canada

I do not have any competing interests to declare.

Adequacy of the overall research design

Review Summary: A randomized controlled design is appropriate to answer the research question. However, the research hypotheses (pages 6 and 7) are not clear. With respect to the intervention, it is uncertain if patients and any accompanying persons will have adequate time to review the intervention, a question prompt sheet (QPS). Several of the outcome measures do not appear to be directly related to the primary study aims or hypotheses. For example, consultation atmosphere (page 15), measured with both the Roter Interaction Analysis System and the Assessment of Interpersonal Motivation in Transcripts is not well-linked to a specific study hypothesis. From a design and analysis perspective, the authors do not appear to take into consideration the potential effect of clustering in their primary analysis even though it appears that patients are nested within oncologists who are nested within centre. Additional details are found in the sections that follow.

R: we have made some changes in the test (highlighted in yellow) and we hope we have addressed the main concerns.

Study Participants

Abstract page 2; Figure 1, page 8; Sample and recruitment, page 11. The protocol does not specify the gender of the patients in the Abstract, in Figure 1, nor in section on Sample and recruitment. In several places, the investigators refer to her or his/her making it difficult to ascertain the gender. Study information found at clinicaltrials.gov indicates that the patients are female; this information should be included in both the abstract and in the section on the sample.

R: we have added "female" where appropriate and checked "his/her" accordingly.

Description of Methods

Recruitment. There is insufficient description of the recruitment process.

R: we have added this information (see page 11)

Adequacy of the Intervention. It is unclear how the QPS intervention is given to the patients and how much time the participants have to review 50 questions. In addition, the QPS is collected by a project staff member prior to the consultation. Presumably this was done so that oncologists would be blind to the intervention or control status of the patient. However, the QPS as originally designed was used by the patient within the consultation to help them ask relevant questions (Butow et al. 1994). The authors should explain how patients will be able to remember all the questions they wanted to ask when the prompt sheet is taken away prior to the consultation.

R: We are aware that in other studies QPS was brought into the consultations, nevertheless we decided to collect the QPS just before the patient enters the consultation, in order to keep the oncologists blind to the intervention or control status. Moreover it was not an easy job to obtain the collaboration to the project from the oncologists. Asking them to change their routine by introducing the QPS into the consultation appeared too much to ask them at that time. Once the study will be completed we will discuss all the procedures and results with the oncologists and we hope that afterwards QPS can be adopted routinely in their practice.

We agree with the reviewer that patients may forget the chosen questions without having a look at it during the consultations. As QPS is collected very shortly before the patients enter the consultation we hope that they can remember some of the most important questions they had in mind. We provided the possibility to check whether during the consultation the patient asked a question from the QPS or the whether oncologist provides that information before the patient asks it. This allows us to see how many of the questions selected on the QPS were either asked or not asked and how many of the questions selected on the QPS were anticipated by the oncologist (and therefore not asked). We think this could be a potential solution to the limitation of the patients of not having the QPS in the consultation and using it as a reminder.

We thank the reviewer for this consideration and we certainly will consider this aspect for the interpretation of the results.

Outcome measures. As stated in the review summary, the authors should better justify the inclusion of outcome measures that are not directly related to the study aims/hypotheses.

With respect to self-reported process and outcome questionnaires, it is possible that patients and any accompanying persons may experience considerable burden in completing so many questionnaires some of which appear to be tangential to the study (e.g., the 24-item subscales of Eysenck Personality Questionnaire).

R: we have made some changes in the test trying to specify the rationale of each questionnaire used. We decided to delete the RIAS as no hypotheses can be made regarding the difference between groups. (see Table 1, pages 12 and 20).

We are aware that patient (and companion) may experience burden in completing so many questionnaires. The overall time required to answer all the questionnaires is about 15-20 minutes for the questionnaires before the consultation and 10-15 for the questionnaires after the consultation (See page 9). Since we have already started the recruitment phase we observed that patients

answered willingly the questionnaires before the consultation because they often have to wait for quite some time before the consultation. After the consultation patients indeed are less willing to collaborate, but questionnaires are shorter. So far we observed about 18% of missing data of questionnaires after the consultation.

Statistical Methods. There are several concerns about the statistical methods. There is insufficient detail to judge the adequacy of the planned analyses. The authors state that they will use t tests in the primary analysis. Do the authors believe that the data will be normally distributed with equal group variances? In one of the references cited by the authors of previous research on the QPS, the data were not normally distributed (Brown et al.). Moreover and as stated in the review summary, the authors do not appear to take into consideration the potential effect of clustering in their primary analysis (Donner and Klar, 2000) even though it appears that patients are nested within oncologists who are nested within centre.

R: Indeed this is a multicenter study. As reported in the paper (see “sample size calculation” section; page 19) a pilot phase was performed: 10 interviews were audio-recorded in each of the three centers. The number of patient questions in these 30 interviews were counted. We found that the data were not normally distributed but the averages and variances were similar in the three centers. The discussion among project members of the three centers excluded differences among centers in the oncologists’ approach to breast cancer consultations. Therefore we choose to not consider the centers as clusters.

Regarding the normality assumption, we followed the suggestion by Lumley et al. (“The importance of the normally assumption in large public health data sets” 2002, Annual Review of Public Health 23: 151-169), who demonstrated the usefulness of the t-test in moderate or large samples also with the non-normal data. However, we will analyze our data also using non-parametric statistical techniques.

In their secondary analysis, they are proposing multi-level modelling which may consider clustering. It is unclear why t-tests are used in the primary analysis and multi-level modeling in the secondary analysis. In addition, the authors state on page 18 lines 27-30 that the primary outcome, significant increase of patient questions, will be compared in the two arms. It is not clear if the authors mean that the primary outcome is a pre-post difference score that will be compared between intervention and control groups or if the primary outcome is the post intervention difference between intervention and controls groups in the total number of questions.

R: The primary outcome is the difference between mean number of questions in the two groups, after the intervention. We will apply the ANOVA approach (specifically t-test for independent samples). Once we will observe that the primary outcome varies in the different subsamples of patient (referred to patients’ or doctors’ baseline characteristics, unknown a-priori), we will perform an ANCOVA using the significant characteristics as covariates.

Once we will observe that the secondary outcomes vary in the different subsamples of oncologists, we will perform multi-level modeling in order to take into account that the patients are nested within doctor.

Standard of written English. On occasion, the authors use imprecise language which does not convey the presumed meaning. One example is the use of the word arguments rather than questions in the instructions to the patients. In many instances, there are typographical errors which are distracting for the reader. For example, see the complete list of authors and Reference 1 in the reference list. There are other issues that are related to writing style. For example, Table 1 does not contain explanations of the short forms so the reader needs to flip back and forth between the Table and the text to find the names of the outcome measures. Table 1 could be improved by including the full name of the outcome measure along with the key constructs measured.

R: we have checked for the writing and made corresponding changes.

Article Summary

There is mismatch between the abstract and the summary with respect to the outcome of patient satisfaction and coping.

R: we add satisfaction and ability to cope also in the abstract (page 2).

Study limitations are not adequately discussed.

The last bullet point in the article summary should be deleted. The QPS as tested in this study is not brought into the consultation.

R: we considered this in fact as a potential limitation of the study; this is the reason why we put this bullet under "strengths and limitation" (page 4). We added a "discussion" paragraph in which we discussed strength and limitations of the study (page 21).

Abstract

Introduction: The authors need to be specific about which aspect of patient involvement they believe is important e.g., in decision making?, in the consultation? The first two sentences are not linked. The second study aim relates to the family member but there are no outcomes listed in the abstract that pertain to this person.

R: we have made changes accordingly to what the reviewer has asked (page 2). Please note also that all the questionnaires used in the study are answered also by the family members.

Methods and Analysis: The primary outcome should be stated. As above, none of the outcomes or the hypotheses is related to the family member. Another relatively minor issue is that the authors use the term family member as well as the terms accompanying person or key person. It would be helpful if the authors could use one term throughout.

R: we have made changes in the abstract (page 2). We decided to use the only term "companion" and change the text accordingly.

References

Two references are missing. The details are included below.

van der Meulen N, Jansen J, van Dulmen S, Bensing J, van Weert J. Interventions to improve recall of medical information in cancer patients: a systematic review of the literature. *Psychooncology*. 2008 Sep;17(9):857-68.

van Weert JC, Jansen J, Spreeuwenberg PM, van Dulmen S, Bensing JM. Effects of communication skills training and a Question Prompt Sheet to improve communication with older cancer patients: a randomized controlled trial. *Crit Rev Oncol Hematol*. 2011 Oct;80(1):145-59. doi: 10.1016/j.critrevonc.2010.10.010. Epub 2010 Nov 13.

R: We cited the two references in the introduction (page 5).

Additional Comments

Page 6 Lines 54 and 57 and Page 7 Lines 3-25. It is quite difficult to assess the logic of this section as it is written. For example, it is not clear what the authors mean by a different perception by the oncologist and by the patient of the doctor-patient relationship. For each hypothesis it would be clearer if the authors included the relevant outcome measure and the comparator. Furthermore, the authors should include more details of how they plan to measure if the question prompt sheet intervention (extended to the accompanying key person) changes the key persons role and participation during the consultation.

R: we have made changes, please see the new text highlighted in yellow (pages 6-7).

Minor Comments

Page 7 Line 52. The authors state that the oncologists complete two measures post consultation but only one is included in Table 1.

R: There was a mistake in the table1 as the CPS is completed by the oncologist after the consultation,

not before. We have changed the table 1 and specified also in the text the name of the two questionnaires.

Page 9 Table 1. CPS. Indicate the number of vignettes and how each is scored. It is unclear when the DP is scored by oncologists. The audiorecording is not a scale but a data collection tool. The OPTION scale does not measure involvement level. It measures health professional behaviours intended to involve patients (Elwyn et al. 2005).

The footnote indicates that certain questionnaires were adapted for companions. The authors need to provide more detail about how this adaptation was done.

R: we provided more details in the measure section (pages 15-18).

Page 11 Line 43. The authors state that the patients and companions complete six questionnaires pre-consultation but Figure 1 indicates three.

R: we added the three missing questionnaires.

Page 11 Line 23. The authors should explain how early stage breast cancer is defined.

R: "early stage" means the absence of metastasis (we added this in the text; see page 11). Presence of metastasis was an exclusion criteria as we suppose possible different information needs related to cancer stage, particularly if the cancer is in a metastatic stage with a poor prognosis.

Page 11 Lines 56-59. The authors state that patients and companions complete six questionnaires post consultation but Figure 1 indicates five questionnaires.

Further they state that projects staff will assist in questionnaire completion. What does assist mean and how will bias be avoided

R: we added a sentence to clarify this point (see page 11).

Page 12 Lines 7-8. The authors need to justify including a questionnaire in which oncologists rate the difficulty of the patient and indicate its relevance to the study aims or hypotheses.

R: we specified better this point in the measures section. We wanted to test whether question asking can influence the oncologist's perception of the patient as difficult patient (page 17).

Further, the authors state on Page 15 that they added an additional four items. What are the psychometric properties of this revised questionnaire?

R: we have added four questions at the end of the questionnaire, leaving the questionnaire as it is (page 17). These questions were:

"How difficult was it to answer to the patient's questions?" (likert scale from 1 to 6);

"How much did the patient seem to be anxious or worried about her health?" (likert scale from 1 to 6);

"In your opinion was the patient depressed?" (yes/ no, if yes: slightly, moderately or seriously);

"Did the patient show to be emotional distressed during the consultation?" (yes /no, if yes: slightly, moderately or seriously).

Page 12 Lines 27,28. Why have the authors indicated the use of Assessment of Interpersonal Motivation in Transcripts in guiding non verbal behaviours during the interaction when the authors are using audiotapes?

R: We changed in the text (page 17). The AIMIT system is applied to dialogue transcripts and differentiates among five motivational systems that guide the verbal and non verbal behaviors during the interactions (attachment, caring, competition, cooperation, seduction). Four of them are reasonably typical of the interactions in medical consultations. Please note that we have replaced the Italian reference to AIMIT with a more recent one in English. (Fassone G, Valcella F, Pallini S, et al. Assessment of Interpersonal Motivation in Transcripts (AIMIT): an inter- and intra-rater reliability study of a new method of detection of interpersonal motivational systems in psychotherapy. Clin Psychol Psychother 2012;19: 224-34)

There is considerable overlap between the text and in Table 1. As suggested previously, an expanded Table 1 would be helpful and would reduce the text description.

R: we have made changes in table 1.

Thank you for asking me to review this manuscript.

References used in this review

Butow PN, Dunn SM, Tattersall MH, Jones QJ. Patient participation in the cancer consultation: evaluation of a question prompt sheet. *Ann Oncol*. 1994;5:199-204.

Brown RF, Butow PN, Dunn SM, et al. Promoting patient participation and shortening cancer consultations: a randomised trial. *Br J Cancer* 2001;8:1273-1279.

Clayton JM, Butow PN, Tattersall MH, et al. Randomized controlled trial of a prompt list to help advanced cancer patients and their caregivers to ask questions about prognosis and end-of-life care. *J Clin Oncol* 2007;25:715-723.

Donner A, Klar N. Design and analysis of cluster randomization trials in health research. New York, New York: Oxford University Press, 2000

Elwyn G, Hutchings H, Edwards A, et al. The OPTION scale: measuring the extent that clinicians involve patients in decision-making tasks. *Health Expectat* 2005;8:34-42.

Reviewer: Phyllis Butow
Professor of Health Psychology,
School of Psychology
University of Sydney, NSW 2006
Australia.

I do not have any competing interests in relation to this protocol.

The protocol lacks a clear set of hypotheses, that address the multitude of assessments being collected.

The statistical analysis section refers to the main study question (do questions increase with the use of a QPS), but do not address many other aspects of the study design.

R: we have made changes specifying better the hypotheses. Please see the new text highlighted in yellow (see pages 13-14).

Further, the comment about interaction analysis involving sequence analysis is not related to the measures being used, and not explicated sufficiently. This may be an aspect of the ROTER analysis, but this needs to be clarified.

R: we have decided to delete the RIAS and the sequence analysis accordingly (see Table 1, pages 12 and 20).

This is an interesting protocol evaluating the impact of a question prompt sheet for breast cancer patients attending their first visit with an oncologist. The authors claim that this is the first such study in Europe, but I am aware of several studies like this which are underway (but not yet published) in Holland, at the AMC.

R: thank you for this information. We will keep an eye in the future on the published results of these researches.

There are many strengths to this protocol, including the careful randomisation and blinding procedures, a clear sample size calculation and careful measurement of outcomes.

I have a few queries about the protocol.

As noted above, I think the protocol needs a series of clear hypotheses to relate the measures to the study questions.

R: we have made changes specifying the hypotheses. Please see the new text highlighted in yellow (see pages 13-14).

The analysis needs to be more detailed to address the hypotheses.

R: The primary outcome is the difference between mean number of questions in the two groups, after the intervention. We will apply the ANOVA approach (specifically t-test for independent samples).

Once we will observe that the primary outcome varies in the different subsamples of patient (referred to patients' or doctors' baseline characteristics, unknown a-priori), we will perform an ANCOVA using the significant characteristics as covariates.

Once we will observe that the secondary outcomes vary in the different subsamples of oncologists, we will perform multi-level modeling in order to take into account that the patients are nested within doctors.

I think the provision of a blank sheet to both control and intervention arms to write down the issues they would like to discuss in the consultation is a good way to control for extra attention, but may reduce the effect size of the intervention compared to control since it will trigger pre-planning in control subjects also. Perhaps the authors need to consider this in the effect size. Have they piloted this approach?

R: We agree on this point. At the time we thought about this and wrote the Italian protocol we discussed in our team three possibilities: 1) to compare QPS with no pre-planning (replication of previous studies); 2) to compare QPS versus no pre-planning, versus a simple open question (three arm design); or 3) to compare QPS+ open question versus a open question only. We have discarded the three arms design as too complicated to conduct from a methodological point of view. We chose the last option in order to test if it is really the QPS to be effective or if a simple preplanning (response to open question) is sufficient. We will of course consider this difference in the interpretation of the results.

Unfortunately we have not piloted this approach and we are in the recruitment phase yet.

(nb issues is a better word than arguments, and should be changed in the protocol. In English, arguments implies that the patient will make a series of points to argue a position.)

R: we have done this.

In the protocol, it suggests that both the blank sheet and the QPS with ticked questions will be collected. Will the patient be left a copy? In previous trials, patients were able to bring the QPS into the consultation to refer to, and it may be less effective if this is not the case here.

R: We are aware that in other studies QPS was brought into the consultations, nevertheless we decided to collect the QPS just before the patient enter the consultation, in order to keep the oncologists blind to the intervention or control status. Moreover it was not an easy job to obtain the collaboration to the project from the oncologists. Asking them to change their routine by introducing the QPS into the consultation appeared too much to ask them at that time. Once the study will be completed we will discuss the results with the oncologists and we hope that afterwards QPS can be adopted routinely in their practice. We added a "discussion" paragraph in which we discussed

strength and limitations of the study (page 21).

We thank the reviewer for this consideration and we certainly will consider this aspect for the interpretation of the results.

Regarding the measures, there was a lot of confusion in the protocol, with different measures being mentioned at different points - these need to be standardised. For example, several measures were not included in the Table of measures, but were referred to later, such as the CPS - pt and the CPS- onc.

R: we have made all requested changes both in the text and in table 1. Please note that now the paragraph on study aims and hypotheses (pages 13-14) precedes the paragraph on Measures (pages 14-19) in order to facilitate the reader.

The main outcome is questions, but it was not clear to me whether this was only patient questions, or companion questions also. I would recommend both, but analysing both separately.

R: Yes, the main outcome was the number of questions asked by the patient. Questions asked by the companion are secondary outcomes and will be analyzed separately.

The questions asked in the consultation will be compared against those ticked prior to the consultation, which is very interesting. However I do not think this comparison can be used to measure unmet information needs, as is suggested in the protocol. The patient or companion may not have asked those questions because the oncologist had already answered them. Unmet information needs could be assessed by the additional questions added to the recall measure (I think), or a much more careful analysis of the information provided and the questions asked would need to be undertaken.

R: We provided the possibility to check whether during the consultation the patient asked a question from the QPS or whether the oncologist provides that information before the patient asks it. This allows us to see how many of the questions selected on the QPS were either asked or not asked and how many of the questions selected on the QPS were anticipated by the oncologist (and therefore not asked). We think this could be a potential solution to the limitation of the patients of not thawing the QPS in the consultation and using it as a reminder.

It was also not clear to me whether companions would be completing demographics, or only the patient. Gender was not listed as one of the variables being elicited from patients and companions, and I would think this was important. If the companion is a key part of the study, it may also be useful to measure the companion's relationship with the patient (e.g. spouse, daughter or son, friend etc) as their role would vary accordingly.

R: socio-demographic characteristics were collected for both patients and companions as well as the type of relationship between them. We have clarified this point in the text (see page 14).

There are many measures and it is not clear to me (in the absence of clear hypotheses) how each is intended to be used.

R: we changed the text in order to make this point clearer (see page 13-18)

Some of the measures I am not familiar with, and it may be important to do more to demonstrate their psychometric properties. The recall questionnaire is based on earlier studies, but does not seem to be related to the methods used in those studies, which were to assess recall by interview with careful prompts.

R: We used validated Italian versions of instruments where available and provided references for the instruments. Validation of the Italian version of English instruments is in progress (DSES, PEI, SDM, SWD, PDRQ-9, DDPQRQ-10, Recall) .

It was also not clear to me how recall would be scored - the authors have added some additional questions related to patients' perceptions of the adequacy of information provision. Will these be scored separately?

R: The Recall Questionnaire allows to count how many and what type of information on treatment the oncologist provided to the patient (analyzing the audio recording) and how many treatment information of what type the patients remembers after the consultations. We are preparing a paper on the recall Questionnaire that describes the instrument in more detail; at the moment we have only preliminary data presented as poster at the 2012 EACH Conference.

The additional questions are independent from the recall questionnaire (see page 16). These questions are:

“Did you succeed in making all the questions you planned to make?”(5 point Likert scale from 1 “not at all” to 5 “very much”).

“In your opinion, how much did the oncologist clearly and completely answer to your questions?”(5 point Likert scale from 1 “not at all” to 5 “very much”).

“How much do you need more information than those received today?”(5 point Likert scale from 1 “not at all” to 5 “very much”).

We will explore answers to those questions separately to the recall questionnaire.

With all the measures, I do not see a measure of satisfaction with the consultation (just with the decision). Perhaps a general satisfaction measure is required?

R: Satisfaction is a complex aspect to explore because highly biased towards positive satisfaction, in particular in critical situations where patients depend on their oncologists . As we have already selected several questionnaires to answer both before and after the consultation we were concerned about the possibility to burden the patient (and the companion), particularly after the consultation. Therefore we decided to concentrate on satisfaction with decision as it was a short questionnaire and more specific for the type of the consultation. In fact, during first consultations oncologists provide information about treatment and usually a decision about treatment is taken. Moreover we suppose that patients will ask more questions about treatment during a first consultation and we want to observe if QPS may have an impact on this specific measure also.

Does the study have adequate power to explore all of these secondary outcomes?

R: Studies in the literature do not provide sufficient information to do this. We will only describe the results on these secondary outcomes and this will be useful for future analysis.

Measures are being collected only immediately after the consultation. Therefore no long-term impacts are being measured. In previous studies anxiety was raised immediately after the consultation but quickly dissipated. Is it worth considering longer term outcomes?

R: This is a very stimulating comment. What kind of longer outcomes you have in mind?

We could think about a follow-up in the future. We welcome any suggestions.

I do wonder whether all the methods of analysing the tapes will be needed - are the interpersonal motivation systems really related to question asking? Again, clear hypotheses would help here.

R: we have made some changes in the text try to clarify better this point.(see page 17)

There will be a LOT of work required to apply all these coding systems to the audio taped consultations.

R: Yes, we agree it is a lot of work. We have started the coding excluding the AIMIT and observed that it takes about 4-5 hours to code a 30 minutes consultation. We think the efforts will be repaid by the satisfaction of the work experience and the final results.

Are oncologists happy to complete forms after every patients is seen? I think there may be a lot of missing data but perhaps the research team is not finding this.

R: The biggest problem with the oncologist was to obtain their consent to audio record their consultation. They were not bothered about the completion of the form after the consultation which requires only a few minutes. So far there are very few missing data.

VERSION 2 – REVIEW

REVIEWER	Mary Ann O'Brien, PhD Assistant Professor, Department of Family and Community Medicine, University of Toronto Canada I do not have any competing interests to declare.
REVIEW RETURNED	31-Jan-2013

THE STUDY	<p>Statistical Methods I do not think it is sufficient to exclude the possible influence of the centre through discussion by team members. The authors have conducted a pilot study so they have some data to better estimate any effect of clustering within oncologist and within centre. With this information, the authors can better plan their analysis. I remain concerned that the authors are proposing to conduct t tests in their primary analysis.</p> <p>Standard of Written English The authors have revised Table 1 to make it clearer to the reader and corrected the spelling errors. Throughout the manuscript, there are several instances where the use of English is not quite correct. For example, in the section on "Secondary outcome measures", the authors use the word "suppose" rather than "hypothesize". The manuscript would benefit from a thorough copy edit.</p> <p>Abstract: Methods and Analysis: This section is more clearly written with the primary outcome stated. A minor issue is the use of the word "supposed" in the following sentence, " All outcome measures are supposed to significantly increase in the intervention group". I suggest the following re-wording "We hypothesize that all outcome measures will significantly increase in the intervention group as compared to the control group".</p>
Additional Comments Overall the authors have made changes to the manuscript that address most of my main concerns. I continue to have some concerns about their analysis (see statistical methods section above).	<p>Additional Comments Overall the authors have made changes to the manuscript that address most of my main concerns. I continue to have some concerns about their analysis (see statistical methods section above).</p> <p>Study Aims and Hypotheses: I am still unclear what the authors mean by "a different perception by the oncologist of the patient's preference regarding her participation....". Similarly, I remain unclear about the meaning of the next line, the authors state " a different</p>

Study Aims and Hypotheses: I am still unclear what the authors mean by “a different perception by the oncologist of the patient’s preference regarding her participation....”. Similarly, I remain unclear about the meaning of the next line, the authors state “ a different perception by the oncologist and by the patient”. The short forms used in this latter statement do not match the short forms in the ‘measures’ section. Could the authors please clarify their meaning?

Procedure , last line.
The authors state the following “...the Assessment of Interpersonal Motivation in Transcripts (AIMIT) [27] which evaluates the five different motivational systems that guide the verbal and non verbal behaviours during interactions”. It is still confusing as the authors are not assessing non verbal behaviours in their study. I suggest that part of the sentence be reworded to something like, “...the Assessment of Interpersonal Motivation in Transcripts (AIMIT) [27] which evaluates the five different motivational systems during interactions”.

Study Participants
It is sufficient to say that the patients are female in the inclusion criteria and Figure 1. It is not necessary to do so throughout the manuscript.

Adequacy of the Intervention.
I understand the authors’ reasons for not allowing the patients to bring the question

perception by the oncologist and by the patient”. The short forms used in this latter statement do not match the short forms in the ‘measures’ section. Could the authors please clarify their meaning?

Procedure , last line.
The authors state the following “...the Assessment of Interpersonal Motivation in Transcripts (AIMIT) [27] which evaluates the five different motivational systems that guide the verbal and non verbal behaviours during interactions”. It is still confusing as the authors are not assessing non verbal behaviours in their study. I suggest that part of the sentence be reworded to something like, “...the Assessment of Interpersonal Motivation in Transcripts (AIMIT) [27] which evaluates the five different motivational systems during interactions”.

Study Participants
It is sufficient to say that the patients are female in the inclusion criteria and Figure 1. It is not necessary to do so throughout the manuscript.

Adequacy of the Intervention.
I understand the authors’ reasons for not allowing the patients to bring the question prompt sheet (QPS) into the consultation. I remain concerned that patients will not remember their questions especially under stress and therefore reduce the effect of the intervention. This issue has now been raised in the “Discussion” section.

Outcome measures.
The authors have better justified the inclusion of so many questionnaires. The 18% missing data post consultation is concerning.

Thank you for asking me to review this manuscript.

<p>prompt sheet (QPS) into the consultation. I remain concerned that patients will not remember their questions especially under stress and therefore reduce the effect of the intervention. This issue has now been raised in the "Discussion" section.</p> <p>Outcome measures. The authors have better justified the inclusion of so many questionnaires. The 18% missing data post consultation is concerning.</p> <p>Thank you for asking me to review this manuscript.</p>	
<p>Additional Comments Overall the authors have made changes to the manuscript that address most of my main concerns. I continue to have some concerns about their analysis (see statistical methods section above).</p> <p>Study Aims and Hypotheses: I am still unclear what the authors mean by "a different perception by the oncologist of the patient's preference regarding her participation....". Similarly, I remain unclear about the meaning of the next line, the authors state " a different perception by the oncologist and by the patient". The short forms used in this latter statement do not match the short forms in the 'measures' section. Could the authors please clarify their meaning?</p> <p>Procedure , last line. The authors state the following "...the Assessment of Interpersonal Motivation in Transcripts (AIMIT) [27] which evaluates the five different motivational systems during interactions".</p> <p>Study Participants It is sufficient to say that the patients are female in the inclusion criteria and Figure 1. It is not necessary to do so throughout the manuscript.</p> <p>Adequacy of the Intervention.</p>	<p>Additional Comments Overall the authors have made changes to the manuscript that address most of my main concerns. I continue to have some concerns about their analysis (see statistical methods section above).</p> <p>Study Aims and Hypotheses: I am still unclear what the authors mean by "a different perception by the oncologist of the patient's preference regarding her participation....". Similarly, I remain unclear about the meaning of the next line, the authors state " a different perception by the oncologist and by the patient". The short forms used in this latter statement do not match the short forms in the 'measures' section. Could the authors please clarify their meaning?</p> <p>Procedure , last line. The authors state the following "...the Assessment of Interpersonal Motivation in Transcripts (AIMIT) [27] which evaluates the five different motivational systems that guide the verbal and non verbal behaviours during interactions". It is still confusing as the authors are not assessing non verbal behaviours in their study. I suggest that part of the sentence be reworded to something like, "...the Assessment of Interpersonal Motivation in Transcripts (AIMIT) [27] which evaluates the five different motivational systems during interactions".</p> <p>Study Participants It is sufficient to say that the patients are female in the inclusion criteria and Figure 1. It is not necessary to do so throughout the manuscript.</p> <p>Adequacy of the Intervention.</p>

Transcripts (AIMIT) [27] which evaluates the five different motivational systems that guide the verbal and non verbal behaviours during interactions". It is still confusing as the authors are not assessing non verbal behaviours in their study. I suggest that part of the sentence be reworded to something like, "...the Assessment of Interpersonal Motivation in Transcripts (AIMIT) [27] which evaluates the five different motivational systems during interactions".

Study Participants

It is sufficient to say that the patients are female in the inclusion criteria and Figure 1. It is not necessary to do so throughout the manuscript.

Adequacy of the Intervention.

I understand the authors' reasons for not allowing the patients to bring the question prompt sheet (QPS) into the consultation. I remain concerned that patients will not remember their questions especially under stress and therefore reduce the effect of the intervention. This issue has now been raised in the "Discussion" section.

Outcome measures.

The authors have better justified the inclusion of so many questionnaires. The 18% missing data post consultation is concerning.

Thank you for asking me to review this manuscript.

I understand the authors' reasons for not allowing the patients to bring the question prompt sheet (QPS) into the consultation. I remain concerned that patients will not remember their questions especially under stress and therefore reduce the effect of the intervention. This issue has now been raised in the "Discussion" section.

Outcome measures.

The authors have better justified the inclusion of so many questionnaires. The 18% missing data post consultation is concerning.

Thank you for asking me to review this manuscript.

VERSION 2 – AUTHOR RESPONSE

Reviewer: Mary Ann O'Brien, PhD
Assistant Professor,
Department of Family and Community Medicine, University of Toronto Canada

I do not have any competing interests to declare.

Statistical Methods

I do not think it is sufficient to exclude the possible influence of the centre through discussion by team members. The authors have conducted a pilot study so they have some data to better estimate any effect of clustering within oncologist and within centre. With this information, the authors can better plan their analysis.

I remain concerned that the authors are proposing to conduct t tests in their primary analysis.

R: We did not find significant differences among the three centres on the pilot sample, we added a sentence on this on pg 19. Our statistician (Co-author Mariangela Mazzi) is confident on the appropriateness of the analyses, which are supported by the data of the pilot study and by the methodological literature cited in the reference list (please see ref 53).

If the reviewer still is concerned about the analyses, Dr. Mazzi is ready to provide further details to the reviewer, may be in a conference call.

Standard of Written English

The authors have revised Table 1 to make it clearer to the reader and corrected the spelling errors. Throughout the manuscript, there are several instances where the use of English is not quite correct. For example, in the section on Secondary outcome measures , the authors use the word suppose rather than hypothesize . The manuscript would benefit from a thorough copy edit.

R: the manuscript has been revised by an English native speaker. Main changes are highlighted in yellow, particularly see pg 5, 15,16,17,20,21.

Abstract: Methods and Analysis:

This section is more clearly written with the primary outcome stated. A minor issue is the use of the word supposed in the following sentence, All outcome measures are supposed to significantly increase in the intervention group . I suggest the following re-wording We hypothesize that all outcome measures will significantly increase in the intervention group as compared to the control group .

R: We have reworded “suppose” with the term “hypothesize”

Additional Comments

Overall the authors have made changes to the manuscript that address most of my main concerns. I continue to have some concerns about their analysis (see statistical methods section above).

Study Aims and Hypotheses: I am still unclear what the authors mean by a different perception by the oncologist of the patient s preference regarding her participation . . Similarly, I remain unclear about the meaning of the next line, the authors state a different perception by the oncologist and by the patient . The short forms used in this latter statement do not match the short forms in the measures section. Could the authors please clarify their meaning?

R: We have clarified this point, please see pg 13

Procedure , last line.

The authors state the following ...the Assessment of Interpersonal Motivation in Transcripts (AIMIT) [27] which evaluates the five different motivational systems that guide the verbal and non verbal behaviours during interactions . It is still confusing as the authors are not assessing non verbal

behaviours in their study. I suggest that part of the sentence be reworded to something like, ...the Assessment of Interpersonal Motivation in Transcripts (AIMIT) [27] which evaluates the five different motivational systems during interactions .

R: We have reworded the sentence as suggested, please see pg 17.

Study Participants

It is sufficient to say that the patients are female in the inclusion criteria and Figure 1. It is not necessary to do so throughout the manuscript.

R: We have deleted the term "female in the redundant places.

Adequacy of the Intervention.

I understand the authors reasons for not allowing the patients to bring the question prompt sheet (QPS) into the consultation. I remain concerned that patients will not remember their questions especially under stress and therefore reduce the effect of the intervention. This issue has now been raised in the Discussion section.

R: We will certainly consider this in the interpretation of the results.

Outcome measures.

The authors have better justified the inclusion of so many questionnaires. The 18% missing data post consultation is concerning.

R: We can be more precise about the missing data at the end of the trial. We make every efforts to have patients completed the after-consultation questionnaires and we are confident that the estimated percentage of missing data of 18% at the end of the trial will be much reduced.