Author’s response to reviews

Title: Pretreatment organ function in patients with advanced head and neck cancer: Clinical outcome measures and Patients’ views

Authors:

Lisette van der Molen (l.vd.molen@nki.nl)
Maya A van Rossum (M.A.van_Rossum@lumc.nl)
Annemieke H Ackerstaff (a.ackerstaff@nki.nl)
Ludi E Smeele (l.smeele@nki.nl)
Coen R.N. Rasch (c.rasch@nki.nl)
Frans J.M. Hilgers (f.hilgers@nki.nl)

Version: 2 Date: 5 August 2009

Author’s response to reviews: see over
Amsterdam, 04-08-2009

Dear Editor,

Thank you for considering our article, “Pretreatment organ function in patients with advanced head and neck cancer: Clinical outcome measures and Patients’ view” for publication in BMC ENT Journal. We appreciate the thorough reviews and the various advices and comments given by the reviewers and we have revised our document accordingly. We have carefully addressed all the issues raised by the reviewers in the order they were presented to us and our reactions and answers are given below.

Yours sincerely,

Lisette v/d Molen
The Netherlands Cancer Institute/ Antoni van Leeuwenhoek Hospital
Plesmanlaan 121
1066 CX Amsterdam
The Netherlands
Telephone: +31(0)20-5122539
E-mail: l.vd.molen@nki.nl

On behalf of the corresponding author:
Prof. Dr. F.J.M. Hilgers
The Netherlands Cancer Institute/ Antoni van Leeuwenhoek Hospital
Plesmanlaan 121
1066 CX Amsterdam
The Netherlands
Telephone: +31(0)20-5122550
E-mail: f.hilgers@nki.nl
Reviewers’ comments:

Reviewer #1:
a. The choice of assessment instruments; this is always a vexing issue on studies such as this as there are no agreed tools to use. However it would seem that the authors did not use any of the validated widely used tools e.g. EORTC. This in itself is disappointing as it decreases the utility of the data they provide…in that clinicians are now used to the common questionnaires and would understand them more than less common ones.

Although this is a justified and understandable comment, there are several reasons why we have decided to ‘only’ include study specific questionnaires and not standardized questionnaires like the EORTC. Firstly, earlier studies in our institute did show that the standardized questionnaires (QLQ-C30 and QLQ-H&N35) were not function specific enough and ‘only’ were useful as a general QOL assessment tool. Study specific questionnaires, on the other hand, gave more useful function specific information, especially when assessing the effects of rehabilitation measures or discreet differences between treatment arms (Op de Coul, B.M.R. et al, 2005; Ackerstaff, A.H. et al, 1998 and 2008). As such, they provide more useful information for clinical practice. Secondly, at this early stage of pretreatment assessment the clinical burden for the patient should be at an acceptable level, and along with the other assessment tools, the questionnaire should be as limited as possible. Therefore it was decided to omit the 65 EORTC questions, which was acceptable for the Protocol Review Board. Thirdly, truly function-specific questionnaires such as the MDADI (Chen, A.Y. et al, 2001) and the Swalqol (Bogaardt, H.C., 2009) were not yet available in Dutch at the start of this research project, thus we decided to use this study specific questionnaire, which has proven its validity in many other studies published from our Institute, as already indicated above. To clarify and explain this in more detail, we added this on the paragraph ‘Assessment aspects – Quality of life, nutrition, pain’, page 4.

b. Furthermore it is not clear of many of the tools used were validated properly previously e.g. FOIS, PAS, Residue scales. Furthermore it appears that the authors used some questions from other tools – in which case the proven validity of the questions is then disrupted. In addition it seems that there were many questions formulated specifically for this study – in which case their validity is not proven. I understand that probably the reason for this is that there were no adequate specific scales, although there are now many scales for assessment of specific functions in head and neck – albeit many have not been validated properly. Therefore this point should be clarified and expanded on in the discussion, and probably included as a limitation of this study.

We can reassure the reviewer that the FOIS and PAS tools used in this study are well-validated and considered to be reliable instruments (see e.g. the publications of Crary M.A. 2005 and Rosenbek J.C. 1996). To clarify this in the article, the terms validated are added when discussing the FOIS and PAS (page 4, lines 16, and 32-33). The other ‘objective’ tools are weight, and mouth opening, and the VAS scale used for measuring pain is the most commonly used measure of pain intensity in pain research (Jensen M.P. et al, 2001; Handbook of Pain Assessment). The use of the residue scales, however, is indeed ‘subjective’ and this is therefore clarified in line 39, page 8. With respect to the validity of the study specific questionnaire, see the answer on the next remark of the reviewer. To clarify and explain the fact that the multidimensional assessment used in this study is not standard we added a new paragraph in the manuscript; ‘limitations of the study’, page 9.

c. The methodology and results on the assessment of validity of the specific questions used for the assessment of specific functions should be clarified in the paper. Again as these are not validated questions (from what I can elucidate) then this should be discussed in the discussion section. For example, it is conceivable that the reason that there is a discrepancy between objective measures and patient subjective reporting is that the questions used are unclear, and not valid! This should be pointed out clearly in the discussion as a potential limitation of the study.

To evaluate the reliability of the questions used in the final analyses, we checked the Chronbach Alpha’s score by Likert Scales. As shown in Table 2, the internal consistency and thus the reliability of
the swallowing questions used is quite good. The questions about the social contacts are slightly less reliable, and therefore are not described as important outcomes. As already mentioned before, we used a questionnaire, which has proven its validity in earlier studies. So, we hope you can understand that we are not fully agreeing with the reviewer’s comment and concern, that the questions used in this study are unclear and not valid. Still it is not a common, standardized questionnaire and therefore, as already indicated above, we point this out in the paragraphs ‘Assessment aspects – Quality of life, nutrition, pain’, page 4 and ‘limitations of this study’ at the end of the discussion, page 9/10.

d. Were the questions given to patients in Dutch or in English? The reason is that some of the questions attached in the appendix are ambiguous e.g. how is your smell? Or, How do you experience your mouth opening? Also the choice ‘rather’. Can I confirm that this is a translation and that the questionnaire was in Dutch?

Yes, the reviewer’s assumption is correct; the questions were given in Dutch. Appendix A gives an overview of the translated Dutch questions and that is why some of them seem ambiguous. To avoid further confusion, this point is added in the title of Table 4, page 17.

Reviewer #2:
a. This study reports function and QOL prior to chemoradiation in patients with stage II and IV cancer. Data was collected it seems as baseline for a RCT and the pre-treatment data is presented in this paper. The most common functional deficits were swallowing, pain and weight loss. These are well known issues and this study is confirmatory. There was not necessarily agreement between the objective and patient derived data. Again this has been shown in numerous papers.

Although not specific questions, we still would like to respond to these remarks. Recently, we have published a systematic review of the research on this particular topic (van der Molen et al. 2009), and to the best of our knowledge, there are no other studies, which compared subjective questionnaires with objectives function assessment tools. Our finding that there was not necessarily agreement between some of the subjective and objective data is not widely published according to our review.

b. In the introduction the authors state that CRT has improved overall survival in all sites. More evidence to support this comment perhaps by giving values by site would be useful.

If possible, that would be an interesting addition, but Pignon et al (2007) only provides overall figures. For the larynx we know e.g. from the paper of Chen et al. (2007) that the outcome for that site is poorer and this reference is therefore added.

c. It is not clear how the need to have comprehensive pre-treatment data influences post-treatment care and outcome. Yes it has value in a research setting but is such assessments really worthwhile in clinical practice. Could the authors strengthen their argument in this respect?

In our opinion, it is of great clinical relevance to evaluate pretreatment functions in this population in order to prevent complications. Patients, who already have swallowing problems or have already suffered from significant weight loss, are at higher risk for developing e.g. aspiration pneumonia or other complications. For those patients it is of great relevance that these problems, either subjectively or objectively, are noted in time, and that preventive measures are taken. So, this is not only a research issue, but should become an evidence-based clinical issue. Furthermore, it is fair to say that the questions asked are quite similar to standard outpatient questions asked during any head and neck consultation, and showing a discrepancy between the answers given by the patients and the outcomes of e.g. videofluoroscopy are certainly clinically relevant.

d. The sample size is relatively small given the length of accrual period. Were they consecutive patients? If not what potential bias might there be. Were there no refusers? Was all data complete?
We are grateful to the reviewer that he has pointed out this omission in our paper, and he is correct in his assumption that not all consecutive patients were included in the study. During the accrual period of roughly 18 months of this clinical trial 72 patients were treated with CRT for advanced head and neck cancer. Seventeen patients could not be included, because of patient refusal (N=4), follow-up known in advance to be abroad (N=2), administrative miss (N=1), cognitive problems (N=6), or physical problems (N=4, i.e. Bechterew’s disease, tetraplegia, jaw problems), leaving 55 patients (76%) for inclusion in the study. This information has been added in the manuscript under Patients and Methods, page 3.

e. A study specific questionnaire was used. It is a great pity that a well-recognized QOL questionnaire was omitted. This would allow the readership to be more familiar with the data and also allow comparison with other studies. Swallowing and pain are both part of standard QOL measures. There are well-validated questionnaires for dietary intake and mouth opening.

These as such valuable and understandable comments are similar to the first two comments of reviewer 1, so we would like to refer to our response to reviewer 1, questions a and b.

f. Weight was assessed from 6 months before treatment but it is not clear how accurate the measurement was at that time. Was it from GP or hospital clinic records?

As is mentioned in the manuscript (page 5), the average weight six months pretreatment was indicated by the patients themselves. So, indeed this measure was not as accurate as a weighing e.g. at the GP’s office, but it is still quite trustworthy, as many people are rather conscious about their body weight. However, this is a limitation, as also was mentioned in the discussion: “These figures have to be interpreted with caution, since they are partly based on the patient’s memory.” (Page 9, line 6)

g. Are the authors advocating that every patient with stage III and IV disease have pre-treatment videofluoroscopy irrespective of stage, site and clinical parameters such as weight loss, low BMI, poor nutritional intake. This is a cost demand on service. Is there no simpler way such as FEES or use of questionnaires to at least identify a more at risk group who would benefit from videofluoroscopy.

Indeed, there are several tools, which could be helpful to visualize pretreatment function problems. We decided to include the most common and validated tools. FEES is a less expensive way to identify swallowing problems, but not all problems will be visualized on this evaluation tool. In the Dutch medical system and more specifically in our institute videofluoroscopy (VFS) is easy to obtain, also since the collaboration with the radiology department is very good. Moreover, we prefer videofluoroscopy, because it allows examination of movement patterns of the bolus and of particular structures in slow motion and frame by frame. VFS studies provide information on bolus transit times, motility problems, and amount, and, most important, etiology of aspiration. The reason why we choose VFS is now described in more detail in the manuscript (page 4, lines 37-40). Still, FEES can be a good alternative, as mentioned in the discussion (page 9).

h. What is the validity of taking number of problems as an outcome?

At this point, the number of problems is merely represented as an observation, and whether it is valid or clinically relevant information remains to be seen at the various post-treatment assessment moments. Reporting on the number of problems will hopefully also create awareness under clinicians, thus enabling them to take into account that the to-be-treated patients may be compromised in terms of different functions even before starting treatment.

i. The study is weakened by the choice of questionnaire and the lack of data of how pre-treatment scores influence post treatment care and outcome. Much of the data is confirmatory to previous literature.
Again, this comment is similar to the first two remarks of reviewer 1 and comment e. of this reviewer, so we hope that the answers given there are sufficient for the first part of the first sentence. The reaction to the second part of this sentence is that we hope to be able to give that answer in due time, when we have those data available from a controlled clinical rehabilitation trial. The expected influence of the pretreatment scores on the posttreatment care is added to the manuscript (page 9, line 25). Finally, we are not as optimistic as the reviewer: there is a surprising lack of data in this respect in the literature, as the already mentioned systematic review we recently published has shown (van der Molen et al. 2009).