Author's response to reviews

Title: Psychometric Properties of two Physical Activity Questionnaires, the AQuAA and the PASE, in Cancer Patients.

Authors:
- Roberto DK Liu (liurob@hotmail.com)
- Laurien M Buffart (l.buffart@vumc.nl)
- Marjolein Spiering (m.spiering@amc.uva.nl)
- Marie Jose Kersten (m.j.kersten@amc.uva.nl)
- Johannes Brug (j.brug@vumc.nl)
- Willem van Mechelen (w.vanmechelen@vumc.nl)
- Mai JM Chinapaw (m.chinapaw@vumc.nl)

Version: 3 Date: 17 February 2011

Author's response to reviews: see over
Dear editor,

We thank you and the reviewers for thoroughly reviewing our manuscript, entitled “Psychometric Properties of two Physical Activity Questionnaires, the AQuAA and the PASE, in Cancer Patients”.

A detailed point-by-point response to the proposed revisions and recommendations by the reviewers is given below. The changes we made in the revised manuscript are indicated with 'tracked changes'/coloured/underlines/highlighted text. We hope that we have satisfactorily addressed the reviewers’ comments and that you will consider our manuscript for publication in BMC Medical Research Methodology.

Kind regards,
On behalf of all co-authors,
Laurien Buffart,
Corresponding author.

Corresponding author’s contact information:
Laurien M Buffart,
EMGO Institute for Health and Care Research,
VU University Medical Center, Department of Epidemiology and Biostatistics
P.O. Box 7057, 1007 MB Amsterdam, The Netherlands.
E-mail: l.buffart@vumc.nl, Phone: +31 (0)20 444 9931, Fax: +31 (0)20 444 8181.
Editorial Request:
- a) supplementary material: you have included your questionnaires as supplementary material, as they have been previously published, please confirm that you have copyright permission to publish them here.

We have decided not to include the supplementary materials in the revised manuscript. In the manuscript, we have included references to the original questionnaires in the methods section (reference number 17 for the AQuAA on Page 6, and reference numbers 18 & 21 for the PASE, also on Page 6).

Reviewer #1:
- Minor Essential Revisions

1. P15 (13) Para 1 – the claim that the construct validity was comparable to other PA self-reports needs some modifications/substantiation, as it is too general as I feel it overstates the situation somewhat, even though on Page 15 (13) Para 1 two recent and comprehensive reviews are cited. I feel both sections somewhat overstate the comparisons between this study and others by claiming others show similar results. A number of currently and widely used PA questionnaires (eg. IPAQ, 7D-PAR, GPAQ), have construct validity often around 0.2-0.3 (Spearman rho, and often statistically significant – although clearly not in all studies – some are indeed poor, but equally some are quite reasonable/fair correlations such as the substantial study by Craig et al 2003, rather than poor correlations). Thus I feel the claim that the low and non-significant rho values seen here (0.05 and 0.16) for the Summary PA are actually quite poor and, arguably closer to the lower end of the spectrum reported and below a typically average/median correlation (if there is such). Hence I think it is a bit unfair to imply these validity correlations are similar (and effectively no worse) than others, when in fact they look lower?

We agree with the reviewer that it may be a bit unfair to state that the validity correlations found in our study are comparable to those found in other studies. According to the Quality Assessment of Physical Activity Questionnaires (QAPAQ) (reference 12: Terwee et al. Sports Med 2010) the quality criteria for adequate construct validity is a correlation with an r≥0.50. In the revised manuscript we reformulated the sentence by removing the word ‘similar’ and we also compared our results with findings from previous studies. See the first and second paragraph on Page 13, under the subheading “Construct Validity”.
2. Table 2: Please define T1 and T0 as well as SDD95 (and arguably SDD95/range95) simply to help some readers so they do not need to have to refer back to the main text and the table can stand alone (not a major problem).

In the revised manuscript, we added the following definitions to Table 2; “T0: the first time the questionnaires were completed, T1: the second time the questionnaires were completed (5 days after the first time), SDD95: smallest detectable difference at a 95% confidence level, range95: the difference between the lowest (2.5\textsuperscript{th} percentile) and highest (97.5\textsuperscript{th} percentile) observed value”. We removed the definition “SDD: smallest detectable difference”. See page 24.

- Discretionary Revisions

3. P5 para 2: perhaps better to state specifically, that “On the 8th day they completed both the AQuAA and PASE at their homes (T0) and after 5 further days the patients completed the questionnaires ....”.

In the revised manuscript we formulated the sentence more specifically: “On the 8\textsuperscript{th} day they completed both the AQuAA and the PASE at their homes (T0). Five days later the questionnaires were completed for a second time (T1)”. See Page 5 under the subheading “Procedures”.

4. Possibly add the Terwee Sp Med review citation (ref 11) on Page 8 Para 3 to support the statement there is no gold standard for measuring PA.

We thank the reviewer for the suggestion to add the publication of Terwee to support the statement that there is no gold standard for measuring PA. We included this reference (12) in the revised version of the manuscript, see Page 9.

5. Possibly mention in the limitations of the Metzger et al MSSE 2008 paper as a source of more recently justified accelerometry cut-points, rather than using the older (but perhaps still more popular) Freedson cut-points.

We thank the reviewer for this useful suggestion. In our study we chose to use the Freedson cut-off score instead of the “newer” Metzger cut-off points. The Freedson cut-off points are generally well-known, and widely used cut-off points, also in studies with cancer patients [13, 14, 24]. This allowed for better comparison of our results with other studies in other cancer populations. We added this information to the discussion, see Page 12.
6. Is the fact that the PASE has such gross “duration” options (1, 2, 3, 4 hours) mean that are almost more nominal than continuous variable, and contribute to its poor correlation?? If so, this seems to be a very unfortunate limitation with this scale and worthy of mentioning in the Discussion? We know most people do not exercise for more than 30 minutes per day – so to have a scale whose smallest unit of duration is “less than one hour” seems very imprecise (people who do 5min and 55 minutes get clustered together and treated as being "the same" - this seems very odd). Surely this scale does not discriminate well among most participants (very large numbers must appear in the “less than 1 hr” and virtually none in the “4 hour” category). Hence it fails at its primary job, to help discriminate durations among a high percentage of the respondents??

It is possible that the choice for the specific duration options of the PASE do not discriminate well among the respondents. This may particularly be the case regarding exercises with short duration (less than 1 hr). However, for other daily activities, such as bicycling/walking or household chores the broader categories may be sufficient to discriminate between active and non-active patients. In the revised manuscript we have addressed this issue as a limitation of the study, see Page 15.

Reviewer #2:
- Minor Essential Revisions

**Background:** I believe the background does not explicitly argue or state WHY it is relevant to be able to measure PA in cancer patients? This may seem very basic, but while anxiety and depression reduced physical fitness and fatigue is mentioned - and the fact that exercise may improve Qol, the rationale behind measuring PA in cancer patients is missing.

Numerous studies have found higher PA levels to be associated with higher QoL and improved survival rates for cancer patients and survivors. Therefore, it is very important to improve PA levels in cancer patients and survivors. To be able to evaluate whether interventions aiming to improve levels of PA are effective, valid and reliable instruments to measure PA are needed. In the revised manuscript, we provided more information of the importance of physical activity with respect to quality of life and survival of cancer patients, see the Background on Page 3.
Background, second last paragraph: The authors state that a PA questionnaire that assesses PA over the past month is less suitable to determine the effect of exercise interventions with a relatively short time frame. In this context it is not clear, why this matters? If the purpose of this study is to identify the best questionnaire for measuring the effect of an exercise intervention - responsiveness of the instrument is relevant and should be tested. Half of the interventions that have been conducted had a duration of 12 weeks or shorter (reference 15: Cramp et al. 2008 & reference 16: De Backer et al. 2009). Therefore, we assumed that assessing PA over the past month might possibly result in less responsiveness to the questionnaire. We applied this as a selection criterion for choosing the PA questionnaires which were to be evaluated. In the revised manuscript we added this information about our selection criteria. See Page 4.

We agree with the reviewer that the responsiveness of a measuring instrument is relevant and that it should be tested. However, the aim of the current study was to evaluate the validity and reliability of the questionnaires. Future studies should evaluate the responsiveness of the questionnaires. This is addressed in the discussion, see page 15. Furthermore, during our study it became clear that recalling PA over the past 7 days resulted in difficulties for most of the participants. With this in mind, recalling over a longer time period (i.e. a month) will be even more troublesome and this will not be beneficial to the responsiveness.

Procedures: Was the AQuAA and the PASE completed at T0 and T1 in a specific or a random sequence? (or ideally: in an randomised alternate sequence?) - please clarify.
The questionnaires were sent to the patients’ homes. We provided no specific directions in which order patients should have filled out the questionnaires. In the revised manuscript, we included the information that no specific directions were given with regard to the order in which the questionnaires should be completed. See page 5, under the subheading “Procedures”.

ActiGraph Accelerometer: I assume the Actigraph accelerometers were uniaxial ones? (‘Vertical accelerations measured by the ActiGraph were converted into activity counts per minute.’) Please include this information explicitly.
For our measurements we used both uni- and triaxial accelerometers. However, for the data analyses we only used the uniaxial data. In the revised manuscript, we mentioned more specifically, that we used the uni-axial data for the validity analyses, see Page 7.
Results/Recruitment and Study Population Characteristics/First sentence: In my opinion this sentence belongs in the Methods section, whereas you may want to add information on how many patients declined to participate and whether any patients dropped out.

In the revised manuscript we provided more information on the number of patients that were invited to participate in the study, and the number of patients who agreed to participate in the method section, under the subheading “study sample”. We found no differences in patient characteristics (age, gender, diagnosis) between responders and non-responders. We added this information to the Methods section on Page 5 and in Table 1.

Discussion/Content Validity/Last sentence: ‘These problems may have introduced recall bias’
- How so - please elaborate?
Due to their disease (and treatment) cancer patients may have a different reference frame compared to non-cancer patients/healthy people. For example because they have a reduced fitness level or may be fatigued, a bias could have been introduced. This became evident during the TSTI (see Table 4), during which a participant experienced making the bed as more intensive than is indicted in the questionnaire. We explained this in more detail in the revised version of the manuscript, see Page 14.

- Discretionary Revisions

Construct Validity: I realise that the terminology in relation to types of validity is indeed debatable, but I would suggest the authors consider calling it 'criterion' or 'criterion-based' validity rather than 'construct' validity. While there is no gold standard for measuring PA, I would argue that you choose to regard accelerometry as a superior measurement or a 'criterion' against which you compare the PASE and the AQuAA.

We agree with the reviewer that this certainly is a matter of debate. In her paper about measurement properties of physical activity questionnaires published in Sports Medicine 2010, Terwee et al. (reference 12) mentioned that there is no gold standard for assessing the validity of PA. Since a true gold standard is lacking, we prefer the term Construct validity, instead of Criterion (-based) Validity.