Author’s response to reviews

Title: Quality of Life in Dementia: a Study on Proxy Bias

Authors:

Alexander M.M. Arons (s.arons@ebh.umcn.nl)
Paul F.M. Krabbe (p.f.m.krabbe@umcg.nl)
Carla J.M. Schölzel-Dorenbos (c.scholzel@slingeland.nl)
Gert Jan van der Wilt (G.vanderWilt@ebh.umcn.nl)
Marcel G.M. Olde Rikkert (M.Olde-Rikkert@ger.umcn.nl)

Version: 3 Date: 5 June 2013

Author’s response to reviews: see over
Dear Arlene Pura,

We would like to thank you for the opportunity to revise our manuscript entitled “Quality of Life in Dementia: a Study on Proxy Bias” (MS: 1353024503822414).

We highly appreciate the reviewers’ constructive and useful comments on the manuscript. We have thoroughly adapted various sections of the manuscript in line with these comments. The adapted sections as well as our responses to the comments are described below in detail.

Thank you for your consideration.

Yours sincerely,

Alexander MM Arons, MSc  
PhD candidate  
Radboud University Medical Centre  
Dept. for Health Evidence  
PO Box 9101  
6500 HB Nijmegen, the Netherlands  
Telephone: +31 (0)243610356  
E-mail: s.arons@ebh.umcn.nl  

Paul F.M. Krabbe, PhD  
Associate Professor  
University of Groningen  
Dept. Epidemiology, unit HTA  
PO Box 30.001  
9700 RB Groningen, the Netherlands  
Telephone: +31-(0)503610896  
E-mail: p.f.m.krabbe@umcg.nl
Response to the comments

Reviewer # 1 (A. Simon Pickard)

Comment 1
“Given the condition of the patients, it seems inconceivable that every single dyad was retained over a one year timeframe. Can you report how many patients dropped out over time? How many caregivers/dyads? Compare characteristics of dropouts after T0 to those retained until T12.”

Response:
The reviewer is correct in pointing out that some patients were lost during the 1 year time period. Table 1 already displayed the different sample constitutions at the three times of measurement. We have added the total sample size at each time point to further clarify this to the readers. Given that Table 1 already presented the absolute numbers as well as the percentages on age, gender, type of dementia, and dementia severity, we believe the readers are well informed regarding the constitution of the sample in this study.

Comment 2
“In the paper, some confusion may arise over terminology as the caregiver self-assessment of HRQL is called proxy-proxy in the paper. This can be confusing, particularly because the framework describing perspectives in ref 38 (Pickard and Knight) use proxy only when a surrogate assessment is made, with “proxy-proxy” describing a proxy assessment of the patient from the proxy’s perspective). Thus, as there is no surrogate assessment, it would be clearer to change the assessment described as “proxy-proxy” to “caregiver self-assessment”.”

Response:
We agree with the reviewer that the terminology might be confusing to some readers. In line with the reviewer’s suggestion we have adapted the term ‘proxy’ into ‘caregiver’ in all circumstances where there are no surrogate assessment implied.

Comment 3
“Methods: is there further detail available on patient and caregiver assessments, e.g. how long did each assessment take (mean, range etc – list in Table 1); what happened if the patient did not understand the question?”

Response:
We have information on the total time it took for the total interviews to complete. However, these interviews consisted of many more instruments therefore we thought these interview durations unsuitable to report. The current manuscript is a methodological sidetrack of the main study (Meeuwsen et al, 2012, Effectiveness of dementia follow-up care by memory clinics or general practitioners: randomised controlled trial), therefore the focus of the interviews wasn’t on the methodological aspects, but on outcomes of HRQoL and costs. If respondents did not understand the question standard protocol was followed (for both the EQ-5D and the QOL-AD). If after a repeated question the respondent still did not understand the question a missing value was recorded. We have added the following sentence to the manuscript in the methods section ‘Respondents’ to
address this comment: "The instrument protocols were followed providing a standardized interview format across respondents."

Comment 4
“Provide guidance for interpreting ICC levels in the methods.”

Response:
We agree with the reviewer that some readers could benefit from extra information regarding the interpretation of ICCs. Therefore we have added the following information to the manuscript: “The strength of agreement between patients and caregivers is expressed as slight (ICC=0.00-0.20), fair (ICC=0.21-0.40), moderate (ICC=0.41-0.60), substantial (ICC=0.61-0.80) and almost perfect (ICC=0.81-1.00).”

Comment 5
“The background notes that very limited research has been done to identify proxy-related factors that explain differences between patient and proxy assessments in dementia. A number of factors have been identified in the broader literature. These are described in reference 38, and would help to inform the discussion section regarding sources of systematic differences/bias."

Response:
We agree with the reviewer and have therefore added the following paragraph to the discussion section: “These newly identified factors differ from those previously identified in other research areas. For example, patient depression has previously been identified as a factor leading to an increase in patient and proxy differences for elderly patients visiting the emergency room [REF]. In addition, burden and psychological distress in caregivers was a significant predictor of patient and proxy differences in psychosocial scores in veterans [REF]. A different study [REF] that focused more on functional status through (instrumental) activities of daily living identified the following factors that contribute to more disagreement between patients and proxies: female proxies, proxies who lived with the patient, proxies who were not first-order relatives of the patient, and proxies who assisted patients with (instrumental) activities of daily living. The newly identified factors thus provide fruitful grounds for new research on systematic differences between patient and proxy assessments.”

Comment 6
“It would be interesting to know if level of cognition is a mediator – was this considered in the analyses?"

Response:
We have no reason to assume that the cognition of caregivers is a mediating variable therefore we did not consider cognition as a mediating variable in the current models. We have explicitly aimed not to include any patient characteristics as we want to study biases that are related to proxies only, not patient-proxy interactions.

Comment 7
“There is nuance to the issue of the validity of proxy assessments. This relates to the inherent problem in defining HRQL in dementia, an important issue raised by the authors in the introduction. If a patient cannot self-report their HRQL or reflect on it, should that rating serve as the gold standard?"
For instance, if a patient is clearly confined to bed, yet they report that they have no problems with mobility, should that disconnect with reality still serve as the gold standard because the patient says it is? For these types of reasons, proxy ratings may be a valid complement to the patient’s rating without agreeing with it. This raises a related issue – the validity of the assessment can depend on the application of the assessment – how will the proxy assessment be used? If for the purpose of substituted judgment, then the patient may be the gold standard, and the desired perspective should be consistent with the application (proxy-patient perspective). If it is to inform clinical decision, e.g. an assessment to inform whether the patient should receive physical therapy, then the proxy-proxy perspective (e.g. a proxy rating of the patient from the caregiver’s perspective) may be important and valid. The purpose of the HRQL assessment receives little mention in the paper and may further inform the discussion section.

Response:
The reviewer raises an important issue here. The distinction between substitution judgments and additional information for clinical decisions is relevant also in the field of dementia. The current study reports on systematic biases on patient-by-proxy assessments. In both contexts will these biases influence patient-by-proxy assessments. Nevertheless, the authors consider this point to worthwhile addressing in the discussion section therefore the following section has been added: “Patient-by-proxy assessments can be used for two distinct purposes. The first is substitution of patient self-assessment. In this situation the patient self-assessment is considered a gold standard to compare patient-by-proxy assessment with. However, when patient-by-proxy assessments are used in addition to patient self-assessment, for example to provide extra information for clinical decision making, then patient-self-assessment should not be considered a gold standard. In this context, proxy reporting might even be more valid than patient-self assessment as the disease progresses. Nevertheless, the biases that were identified might occur in both substitution judgments and informing clinical decision making.”

Comment 8
“I am aware of some evidence to support your findings on projection bias in stroke. Pain and depression in caregivers systematically biased assessments, which could be considered a form of “projection bias” (Hung SY et al, 2006, J Clin Epid).”

Response:
We find this reference a worthwhile addition to the background and have incorporated it in the Introduction section.
Reviewer # 2 (Hilde Verbeek)

Comment 1
“Why did the authors not include the stage of dementia in the linear mixed models, as it is known that this has an important impact on the ratings of QoL for people with dementia? It could be that for people with dementia who experienced more severe cognitive impairments, it is more difficult for proxies to assess their quality of life compared with people with less cognitive impairments. In my opinion, the authors should take this into account.”

Response:
The aim of the current study was to assess which caregiver characteristics systematically influence patient-proxy HRQoL assessments. As such, we have not included any patient characteristics in the model. We agree with the reviewer the stage of dementia will probably have a significant impact on the HRQoL assessments. However, this is not a bias that we are interested in in the current study. Following the reviewers line of thought other characteristics such as patient cognitive functioning, level of depression, GDS stage, social functioning status, level of agitation, word finding problems, etc. could all make it more difficult for proxies to assess the HRQoL of patients. However, these are biases that are caused by patient-proxy interactions, not the proxy him/her self. In this article we are solely interested in proxy ‘projection’. To make the above line of reasoning more explicit in the manuscript we have rephrased the aim of the study into: “The aim of this explorative study is to assess whether certain caregiver characteristics contribute to a bias of patient-proxy HRQoL assessments.”

Comment 2
“Please add the number of people included in the analyses. Were QoL AD scores and EQ5D scores available for all 175 dyads of people with dementia and their proxy? How many dyads were included in the linear mixed model analyses? Usually, the completion rate of QoL-AD is lower for people with dementia compared with proxy ratings.”

Response:
Firstly, we would like to point out to the reviewer that in these analyses only data from caregivers was used. Patients were not included in the linear mixed model because the research question we wanted to address was whether or not proxies project part of their own quality of life onto patients. Therefore, we used proxy characteristics are predictor variables on proxy assessments of patient quality of life.
The following sentences have been added to the results section (for VAS): “This model had a combined 498 observations across the three measurement times. A total of 407 observations were used in the model.” and (for utility): “This model had a combined 525 observations across the three measurement times. A total of 486 observations were used in the model.”

Comment 3
“QoL-AD The authors use the QoL-AD to assess both patient and proxy quality of life. To assess proxy quality of life, this is an uncommon procedure. It would be helpful for readers if the authors included a reference of a study in which this procedure has been validated or used, as the original reference the authors mention (39), does not address this issue.”
Response:
We have added a reference as per the reviewer’s request where the validation of the caregiver’s own QoL is addressed (“Validation of the Brazilian version of the quality of life scale for patients with Alzheimer’s disease and their caregivers (QOL-AD”).

Comment 4
“Please specify in the method section whether the EQ5D was used to measure proxy QoL or also the patient perspective. This is currently unclear for readers.”
Response:
The methods section has been expanded. The section regarding the EQ-5D now contains the following additional information: “The EQ-5D was used for patient self-assessment, patient-by-proxy assessment and caregiver self-assessment.”

Comment 5
“Analyses. It would be interesting to also analyse the data using the patient QoL-AD score as a dependent variable and the proxy QoL AD score as an independent variable. Why did the authors choose only to analyse models using the EQD as a dependent variable?”
Response:
The aim of this explorative study is to assess whether certain caregiver characteristics contribute to a bias of patient-by-proxy HRQoL assessments. Therefore, if we would have used the patient’s self-assessed QOL-AD scores as a dependent variable, then we are no longer investigating projection biases, but addressing a different research question, one regarding agreement/concordance. We certainly agree with the reviewer that this would be interesting, but it would also be out of the scope of the current manuscript. Additionally, those types of study designs have been reported far more often than the current study design, which focuses only on proxy projection.
The reason why we chose for a different instrument is because of the possible overlap QOL-AD items. A good example is the item ‘financial situation’. In our study, the majority of dyads were partners, which in the Netherlands means it is very likely that financial situation is identical for both members of the dyad. This type of overlap in item responses obscures potential biases, and will only increase the amount of ‘noise’ in the model.

Comment 6
Could the authors please add the time points for the statistical models data? I.e. please specify at which time points the EQ5D and QoL-AD scores were assessed for use in the model (e.g. only baseline data?) Furthermore, it would be helpful if the authors added how they dealt with missing data.
Response:
We have added: “These variables were used in the model at T=0, T=6 and T=12.” To the analyses section. The missing data were excluded from analyses list wise (as is always the case in linear mixed models). The data structure for the linear mixed model allows for better missing data handling than a repeated measures ANOVA, which is why we chose this type of model.
Comment 7
“The authors should state in the results section that the overall explained variance of the model is relatively low (R²=0.095)”

Response:
The R² is indeed low, but this is a desirable outcome. Ideally, we would wish for a situation in which all of the variance in patient-by-proxy assessments would be explained by characteristics of the patient, and none by proxy characteristics (R²=0.00). The current manuscript indicates that (at least) 9.5% of the variance is a result of proxy projection. The reviewer suggests mentioning the R² in the results section and indicating that it is low. We prefer refraining from writing subjective interpretations in the results section. However, we agree that the R² deserves more attention, therefore we have added the following paragraph in the discussion section:
“It should be noted that the explained variance of both linear mixed models was low. However, with the current study design, this is a desirable outcome. If the models would explain all of the variance then this would imply that patient-by-proxy assessments would only be based on proxy characteristics and not on patient characteristics. In this study the explained variance in the models was less than 10% which suggests that the bias that is present in patient-by-proxy HRQoL assessments is small compared to the influence of actual patient characteristics.”

Comment 8
“Why do the authors use a numbering for the variables related to proxy bias and start with age, the least strong characteristics? It appears a bit odd to read that first and least strong predictor was age. On what is this numbering then based on? I would suggest to start with the most important one.”

Response:
We have addressed this issue by reordering the three paragraphs that deal with biases on the VAS scale.

Comment 8
“The authors’ explanation for how ‘money’ has an influence is highly speculative in my opinion and not supported by any literature. The authors should compare this with other known literature or delete this from the manuscript, in my opinion, as it appears a too arbitrary judgement. This also accounts for the reasoning why ‘ability to do things for fun’ was related. Please delete or compare it with other known literature. For example, an alternative explanation for the relation could be that proxies who do more things for fun experience less burden and people who experience less burden may rate QoL of their relative in a more positive way than people who feel burdened.”

Response:
We agree with the reviewer that our interpretation of how money biases VAS ratings is speculative. Indeed, in the manuscript we state: “A possible explanation for this relationship could be that financial status functions as a mediating variable for socio-economic status (SES).” These factors are not directly related to ‘ability to do things for fun’ as this variable was only significant in the model on EQ-5D utility and not on the VAS ratings. Furthermore, we have compared our findings to other literature. However, we have not been able to identify previous literature that investigates this very issue. That is why we state: “This is a new finding and to the authors’ knowledge has not been
previously reported elsewhere.” We believe that this finding is important and that other researchers might be stimulated to address this issue in future research. Therefore, we believe that this issue is addressed appropriately.