PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Validity of the PROMIS-10 Global Health assessed by telephone and on paper in minor stroke and transient ischemic attack in the Netherlands
AUTHORS	Lam, Ka-Hoo; Kwa, Vincent

VERSION 1 – REVIEW

REVIEWER	Dipankar Dutta Stroke Service, Gloucestershire Royal Hospital, United Kingdom
REVIEW RETURNED	10-Dec-2017

GENERAL COMMENTS	This paper describes the use of the Dutch version of the Patient Reported Outcomes Measurement Information System 10-Question Short Form (PROMIS-10) in a single centre cohort of 75 minor stroke or TIA patients who had their index event a year previously. The PROMIS-10 was completed by patients or carers on paper and by telephone during the second half of the study. The Dutch version of the SF-36 was used as a comparator. The study should be of some interest but I have identified a few problems which need addressing: Abstract – The introduction/objectives (lines 5-17) suffers from some typos and incomplete sentences. Some of the introductory sentences do not add value to the paper and could be replaced by more relevant statements. Introduction – the whole of the first paragraph (lines 15-27) is redundant and could easily be deleted. Sentences like (line 25) "Patients usually receive further rehabilitation treatment after suffering a major stroke" do not add anything to the paper. Instead of this digression, I think the authors should have expanded on why patient reported outcome measures are need in TIA and minor stroke and explain the background and rationale for this study in greater detail. A little bit of background on patient populations on whom the PROMIS-10 has been validated previously would be useful to provide justification for this study. It would be useful to know if the PROMIS-10 has been studied in minor stroke or TIA previously. Methods –Line 23 -It is not clear what the initial verbal consent applies to. Did the patients consent only to receive the study information or did they consent to their data being accessed? This should be made clearer. A comment on why the authors decided to use the outcome measures at one year would also be useful. Subjects- 45-47- a comment on whether MRI was used in diagnosis would be useful. The authors state that patients not referred to the BAC (stroke advice centre) were excluded. Although they acknowledge this in the limitations, further information on why this was the case and selection criteria for BAC referral would be useful. Data analysis –The Pearson's correlation coefficient (r) is used in

	this paper although this has been criticised as inappropriate by many statisticians. Other methods such as the Bland-Altman method or some modification of the correlation coefficient are recommended. However, the correlation coefficient is widely used in the literature in many peer reviewed journals. The authors should consider reanalysing their data using other methods or provide a justification for the use of the correlation coefficient and state if other methods were considered. Results – A study flow diagram would be useful to explain the exclusions which left just 75 patients included in the study. Table1 provides data on the study population (n=75) and non-responders (n=182). Where did the data for the non- responders come from? How did they consent to their data being used? Was the consent verbal? Much more detail is needed on the exact consent process as not responding may be taken to mean lack of implied consent. In results lines 23-27, it is stated " ..182 were non-respondents (108 had no baseline measurements, 12 were insufficient proficient in Dutch,11 died before follow-up at one year post-stroke, 6 had dementia or a behavioral disorder, 28 refused participation, 13 were not responsive, and 4 were not reachable by phone)." The authors will need to explain in greater detail how data from the "non- responders" is being presented. The discussion is reasonable and limitations are acknowledged. Finally, some bits of the paper need revision for language as some errors have crept in. e.g. abstract line 11, introduction line 21, 25-29, 41, 53, 54. Methods 31-33 to name a few.

REVIEWER	Maggie Lawrence Glasgow Caledonian University, Scotland, UK
REVIEW RETURNED	14-Jan-2018

GENERAL COMMENTS	The paper is of interest to health professionals working with people diagnosed with TIA. However, I would suggest that some changes are required, as described below Generic comments More detail and clarity required in the methods section Stronger 'arguments' need to be made in the Discussion Proof-reading for use of English, including tense is recommended as there are minor errors throughout the paper The tables require formatting, including insertion of (SD) and %, as appropriate Minor changes p.4, line 19: QoL should be written in full here p.4, l. 25: RAND-36 is introduced – a brief explanation should be provided p.4, l. 27: Specify dates rather than 'first half' p.4, l. 31: proxies are mentioned – was data collected reading this or is the number of proxy completions unknown? pp. 4-5, l. 55 on: I think there is a typo in this sentence – one too many 'not's perhaps? p.6, l.21: the numbers don't add up - by my calculation there would have been 267 eligible patients Table 1: are the authors satisfied that the groups are sufficiently similar to not warrant comment? For example, differences between the telephone group and the non-respondents in terms of gender, TIA/stroke, incidence p.10, l.40: the argument made for fluctuations in physical health as opposed to mental health needs to be strengthened and no mention

	is made here of the time lapse between completion of the two instruments p.10: l. 50 and 56 seem contradictory p.11: some evidence from research literature is required to support the statements made about telephone completion of outcome measures p.12 l. 21: PROMIS-10 was designed to be used at 3-months post-stroke – no rationale is provided for its use in this study at one-year (and this should provided in the Methods section)

REVIEWER	Dr Louise Marston, Principal Research Statistician UCL, London
REVIEW RETURNED	01-Mar-2018

GENERAL COMMENTS	Introduction - Express incidence per 1000 or 10000 inhabitants to allow whole numbers as they are easier for the reader. Data analysis p5 - You do not say what the t-tests and chi 2 tests are used for, please add more detail. Table 1 - spell table correctly in the title. Table 1 - describe what low, medium and high is in terms of education. Table 1 - much better to show exact p-values if you are going to do statistical tests so the reader can make their own judgement. P>0.05 is generally discouraged. Construct validity - I suspect correlations differ between data collection methods because of random variation, partly brought about by the sample size being small. This has been acknowledged in the discussion. In terms of inter method reliability, most differences between the two data collection methods are not statistically significant, but what about clinically significant? I would have doubts about inter method reliability if there was a clinically significant difference. Figure 2 - Boxplots should show median (IQR). These do not fit with Table 3 which shows mean (SD), so I do not think Figure 2 is adding anything useful. Add the limitation that you did not do test-retest analysis because data were only collected from everyone at one time point.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1:

- Abstract:

Spelling errors and incomplete sentences have been corrected, redundant sentences have been removed or replaced.

- Introduction:

The first paragraph has been shortened and made more relevant with regard to the rationale of the paper. The extent in which Dutch PROMIS-10 has been studied in TIA and minor stroke has been added.

- Methods:

The sentence concerning verbal consent has been clarified. The rationale for assessment of health at

one year post-stroke has been added.
- Subjects:
The use of MRI for diagnosis has been clarified.
BAC assessed health at baseline (shortly after discharge), in a concurrent study at OLVG health is measured again in the same population at one year post-stroke with the purpose of comparing health outcomes at both timepoints. As the patients are subsequently included for the current study, patients without baseline health measurements are not included.
- The appropriate and inappropriate use of the correlation coefficient has been acknowledged. Bland-Altman plots have been constructed and added to the manuscript to complement the correlation coefficient.
- Results
A study flow chart (figure 1) has been added.
The data for the “non-respondents” (non-responders and excluded patients, as these were all patients considered eligible) are collected through medical records. This was done without the need for consent of the patients, since efforts of acquiring consent was not in proportion to the data to be collected, in addition to the following circumstances (Code of Conduct for the Use of Data in Health Research 2004):
o Data collection for scientific research purpose;
o No objection was present, known, or expressed to this manner of data collection;
o Only non-identifiable information was collected, and the data was collected in a unidentifiable way (anonymously).
The data from the “non-respondents” is clarified in the Methods section.

Reviewer 2:
- The methods section has been more detailed and clarified.
- The arguments in the discussion section have been expanded and strengthened, especially the difference between physical and mental health, and the difference between paper and telephone assessment.
- In the caption of table 1 it is noted that all data are expressed as n (%), except where specified (which is only the case for follow-up duration and age).
- The minor changes have been acknowledged and adopted in the manuscript:
o Regarding the mentioning of proxies: in our study the study documents were intended to be completed by the subject on their own. However, of at least one occasion a subject reported to have completed the study documents with help of a proxy. Hence, the mentioning of “or with help of a proxy”. Documentation of proxy completions is unknown, except for the telephone assessed PROMIS-10.
o Regarding the number of eligible patients: we confusingly mentioned the 291 patients who were referred to the BAC (= eligble), when the 301 who were not referred to the BAC (= non-eligible) should have been mentioned. This has now been correctly worded and a flow chart (figure 1) has been added to further clarify the study population.

o There were no statistically significant differences between the telephone and paper assessment group. For the characteristics (gender and diagnosis) that did show obvious differences (although not statistically significant) subgroup analysis were performed and mentioned in the discussion. for gender, and diagnosis (TIA or minor stroke).
o Literature relevant to the statements made about telephone assessment in the discussion has been added.
o The rationale for the assessment of PROMIS-10 at one year has been provided in the Methods section.

Reviewer 3:
- The use of the t-tests and chi square tests are clarified in the methods sections and also in the note

under table 1.
- Low, medium and high education has been clarified in the methods section.
- Exact p-values have been added to table 1.
- Regarding the comment about inter-method reliability: strictly speaking, no data has been gathered to make conclusions regarding clinically significance. However, the observed differences, including the statistically significant difference, between assessment methods are relatively low (within one standard deviation). Test-retest reliability along with analysis for responsiveness of the Dutch PROMIS-10 would have been of interest but unfortunately has not been carried out.
- Figure 2 has been left out.
- The lack of test-retest analysis has been added.

VERSION 2 – REVIEW

REVIEWER	Dr Louise Marston, Principal Research Statistician UCL, London
REVIEW RETURNED	20-Apr-2018

GENERAL COMMENTS	Table 1 - final column - I assume the p-value is non-responders versus all responders? Table 1 - localization for responders versus non-responders should be analysed using Fisher's Exact test not chi squared due to small numbers in a number of cells (this is probably the case for marital status between phone and paper). Figures appear to have not reproduced in the pdf so was unable to comment on these. More explanation on the interpretation of the Bland-Altman test is needed. Please cite Bland and Altman's Lancet paper on the Bland Altman method. Table 2 - check 0.88 (95% CI 0.78, 0.90) is correct. Lines 35 to 39 are difficult to understand and may not be results of interest given the correlations between methods look similar, and any differences are likely to be because of the small numbers and think you are overstating the differences.

REVIEWER	Dipankar Dutta Gloucestershire Royal Hospital, UK
REVIEW RETURNED	23-Apr-2018

GENERAL COMMENTS	Thank you for improving the paper in keeping with editorial and reviewer comments. All my concerns have been adequately addressed.

VERSION 2 – AUTHOR RESPONSE

Reviewer 3:

- Table 1: the p-values in the final column are indeed p-values for the differences between the respondents (i.e. study population, n = 75) and non-respondents (n = 182). Where Chi Square tests were used inappropriately, the Fisher's Exact test has now been used.

- Bland-Altman test: the methods and interpretation of the Bland-Altman plots have been expanded in the results and discussion sections. Bland and Altman's Lancet paper has been cited.

- Table 2 has been checked. Correlation coefficient and confidence interval between RAND-36 and PROMIS-10 physical health (on paper) has been corrected to 0.88 (95% CI 0.78, 0.93) from 0.88 (95% CI 0.78, 0.90).

- Lines 35 to 39 are largely rephrased or removed. Differences most likely due to small sample size has been acknowledged.