

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Speech and communication in Parkinson's disease: a cross-sectional exploratory study in the United Kingdom
AUTHORS	Barnish, Maxwell; Horton, Simon; Butterfint, Zoe; Clark, Allan; Atkinson, Rachel; Deane, Katherine

VERSION 1 - REVIEW

REVIEWER	Laura Fogg-Rogers Research Fellow University of the West of England, Bristol UK
REVIEW RETURNED	18-Oct-2016

GENERAL COMMENTS	<p>This is an interesting and well designed study to address gaps in the literature around speech intelligibility and functional communication in Parkinson's disease. The literature and need for the study are well laid out, along with the design and measures used to address these objectives. The results and statistics are clearly presented with well defined key findings in the discussion. The study highlights the need for more functional communication therapies.</p> <p>Please address some of the grammar and flow in the Introduction third paragraph - some sentences stop halfway through and are then completed in the following sentence.</p> <p>Three times in the paper (introduction and first/last sentence of discussion) a reference is made to a presumed Theory of Change for links between cognitive status and functional communication in PD. This is described as a pathway for how these issues may occur. Please expand on this pathway to articulate exactly how you think limitations in functional communication may occur. The reader can guess at this, but it would be nice to have it fully explained.</p> <p>One finding is that intensity relates to intelligibility, and also that female PD participants have lower intensity. Please expand on gender effects and whether female PD participants were generally less intelligible and if that relates to functional communication.</p> <p>While it is stressed that intelligibility does not fully explain functional communication, few clues are offered as to why this may be. Some more explanation on other factors for effective functional communication could be useful - e.g. prosody etc. As it is stressed that functional communication should be worked on in therapy and not just intelligibility, some ideas for areas to develop may be useful for SLTs.</p>
-------------------------	--

REVIEWER	Stacey Humphries Postdoctoral Scholar University of Pennsylvania USA
REVIEW RETURNED	31-Oct-2016

GENERAL COMMENTS	<p>This study provides an important contribution to our understanding of how functional communication is affected in Parkinson's disease, but does need revision. The manuscript would benefit considerably from an explicit description of the research questions addressed and hypotheses tested. At present, given the huge array of tests (uncorrected), it is not clear to the reader what exactly is being examined and why. At times, key methodological decisions appear to have been made arbitrarily so please justify these clearly. Grammatical errors in the writing are common and the correction of these would substantially improve readability. Please address the specific comments below.</p> <p>Abstract – needs greater clarity: “associations” – between what? Not clear “various stages along the potential pathway” – what are the stages? How was this defined in the cross sectional sample? “cross sectional study” – of what? Of patient voice acoustics? Functional communication? Results: This is the first time we're hearing of the distinction between “read” and conversational” sentences so this is hard to follow. ...”and this was associated with worse cognitive status” – presumably of the speaker not the listener.</p> <p>Introduction: “Speech and functional communication difficulties are also widespread.” – It would be worth outlining here the distinction between motoric speech impairment and cognitive language impairment. “iii) how closely related difficulties” IN “speech”? “iv) also how wider aspects of communication such as emotional conveyance may be affected.”- This seems to come out of nowhere and should be justified earlier. Why did you choose to look at emotional conveyance specifically? “and also be an important predictor” – change: and is an important predictor Sentence beginning: “A systematic review...” is grammatically flawed. Revise. “None of the studies had used... a validated outcome measure that assessed either communicative effectiveness of communicative participation as a concept, rather than specific sub aspects...” Why is this important? Needs further justification. Does ICF need to be defined? Remove: ...”the practicalities of such a large sample size probably explain why”... “In the only study to assess normal listeners’ ability to identify specific emotions in the speech of people with PD...” Please also discuss here Pell, Cheang & Leonard, 2006. At the end of the introduction please provide an in-depth overview of your exact research questions and hypotheses.</p> <p>Method: First section: please say more here about what the cross sectional design was actually assessing.</p>
-------------------------	--

	<p>Participants: How many PD participants and CPs were recruited?? Critical - How was the sample of 20 PD and 20 CP chosen from the larger sample? “Since CPIB was our primary measure...” Why? Justify earlier. What were the 3 sentences from Miller and the one additional sentence read by the participants? Was the reliability check conducted on 10% of the data from every participant, or 100% of the data from 10% of the participants? If the latter, this could be problematic since the acoustics of people with Parkinson’s will vary considerably from person to person. Discuss. “In the intelligibility task, all stimuli were presented audiovisually, while in the emotional conveyance tasks, half were presented audiovisually and half in audio only.” Why? Justify this decision. On what basis was the predicted effect size $r=.05$ chosen for use in the power analysis? Was this based on the result of previous studies? Please cite them here. Please report the results of the test-retest reliability and convergent validity of CPIB in the method not the results. How was motor impairment assessed? UPDRS, Hoehn and Yahr etc. Were transcribers allowed to listen to the sentence only once or more than that? If once, were the sentences of the two groups matched in length?</p> <p>Results: Why is smoking status described? Please describe earlier why this information was collected and what it tells us. Please include the exact test statistic, p value, effect size and 95% confidence interval for all between group comparisons reported in the results. Was there an association between % words correctly transcribed, and emotional conveyance?</p> <p>Discussion: In the emotional conveyance section please also discuss your findings with relation to Pell, Cheang and Leonard, 2006. Why would the happy emotion be particularly affected? Many of your findings were more strongly associated with gender than with PD. Please discuss. “Intelligibility did not account for a large proportion of variance in functional outcomes” – was it not the case that intelligibility predicted both communicative participation and effectiveness? As far as I can tell, level of motor disability was not assessed. Since greater cognitive impairment is likely to be strongly correlated with greater motor impairment (as both decline in parallel), can you truly say that cognitive factors are more important than motor factors (i.e. could there be a confound here)? Despite the fact that there were very few differences between the two groups on your acoustic measures, it was still the case that patients were significantly less intelligible. This indicates that motor speech impairments still exerted a strong general effect, even though this was not captured in the fine grained analysis of different types of speech acoustics. Could it have made for a stronger study design to assess motor disability more directly to rule this out? Please discuss.</p>
--	---

VERSION 1 – AUTHOR RESPONSE

Reviewer Name: Laura Fogg-Rogers

Institution and Country: Research Fellow, University of the West of England, Bristol, UK Competing

Interests: None declared

Comment: This is an interesting and well designed study to address gaps in the literature around speech intelligibility and functional communication in Parkinson's disease.

Response: We thank the reviewer for her positive appraisal of our study.

Comment: The literature and need for the study are well laid out, along with the design and measures used to address these objectives.

Response: We thank the reviewer for this positive comment.

Comment: The results and statistics are clearly presented with well defined key findings in the discussion. The study highlights the need for more functional communication therapies.

Response: We thank the reviewer for this positive comment.

Comment: Please address some of the grammar and flow in the Introduction third paragraph - some sentences stop halfway through and are then completed in the following sentence.

Response: We have revised this paragraph to improve its expression. See tracked changes.

Comment: Three times in the paper (introduction and first/last sentence of discussion) a reference is made to a presumed Theory of Change for links between cognitive status and functional communication in PD. This is described as a pathway for how these issues may occur. Please expand on this pathway to articulate exactly how you think limitations in functional communication may occur. The reader can guess at this, but it would be nice to have it fully explained.

Response: We have thoroughly considered your comment and how best to address it. We are restricted with regard to the length of our manuscript. We have therefore added the following text at the end of the discussion as a take home message: "The pathway to functional communication difficulties in PD is likely to involve complex, multi-factorial mechanisms for change, including for example motoric, cognitive and psychosocial elements. Future confirmatory research should aim to clarify the elements and mechanisms of this pathway, as well as how they may differ between individuals with PD, which is a condition known to vary considerably in its clinical expression" In addition, we have added the phrase "elements and potential mechanisms for change" to the introduction.

We have added the following reference:

- Walters EC. Variability in the clinical expression of Parkinson's disease. J Neurol Sci 2008; 266: 197-203.

Comment: One finding is that intensity relates to intelligibility, and also that female PD participants have lower intensity. Please expand on gender effects and whether female PD participants were generally less intelligible and if that relates to functional communication.

Response: Gender was not a statistically significant predictor of intelligibility. We have added a sentence to the results section to say this

Comment: While it is stressed that intelligibility does not fully explain functional communication, few clues are offered as to why this may be. Some more explanation on other factors for effective functional communication could be useful - e.g. prosody etc. As it is stressed that functional communication should be worked on in therapy and not just intelligibility, some ideas for areas to develop may be useful for SLTs.

Response: We have ensured that it is now clearly stated in the discussion that future confirmatory research is required to confirm and fully explain the pathways to functional communication difficulties in PD. With regard to areas for clinical development, we have added the following text: "In achieving this, it is important to consider what the particular client's communication needs and goals are, what challenges the client faces in accomplishing these, and what approaches may facilitate this. It is important to remember that communication needs differ between clients, and that clients differ in what they consider full participation in life."

Reviewer: 2

Reviewer Name: Stacey Humphries

Institution and Country: Postdoctoral Scholar, University of Pennsylvania, USA Competing Interests: None declared

Comment: This study provides an important contribution to our understanding of how functional communication is affected in Parkinson's disease, but does need revision.

Response: We thank the reviewer for this positive comment.

Comment: The manuscript would benefit considerably from an explicit description of the research questions addressed and hypotheses tested.

Response: We have re-written the end section of the introduction in which we state our aims. Please see tracked changes. Our study is exploratory and is framed in terms of research questions rather than hypotheses. We have also added explanation with regard to why functional communication was our primary outcome.

Comment: At present, given the huge array of tests (uncorrected), it is not clear to the reader what exactly is being examined and why. At times, key methodological decisions appear to have been made arbitrarily so please justify these clearly.

Response: We have addressed your specific comments as below

Comment: Grammatical errors in the writing are common and the correction of these would substantially improve readability.

Response: We have addressed issues of writing style and grammar. See tracked changes.

Comment: Please address the specific comments below.

Response: See below

Comment: Abstract – needs greater clarity:

"associations" – between what? Not clear "various stages along the potential pathway" – what are the

stages? How was this defined in the cross sectional sample?

Response: While adhering to the structure and length requirements of a BMJ Open abstract, and in particular, the purpose of this 'objectives' section, we have rephrased this to clarify. See tracked changes.

Comment: "cross sectional study" – of what? Of patient voice acoustics? Functional communication?

Response: We have re-phrased the 'design' section of the abstract to clarify as much as possible within the constraints of the abstract length and structure.

Comment: Results: This is the first time we're hearing of the distinction between "read" and conversational" sentences so this is hard to follow.

Response: We have added "on read and conversational speech" to the 'Participants' section of the Abstract, which precedes this, so should mean that this no longer comes so suddenly in the results of the Abstract

Comment: ..."and this was associated with worse cognitive status" – presumably of the speaker not the listener.

Response: We have added the word "speaker" into this sentence.

Introduction:

Comment: "Speech and functional communication difficulties are also widespread." – It would be worth outlining here the distinction between motoric speech impairment and cognitive language impairment.

Response: We have now introduced the ICF framework at this point of the manuscript in order to clarify the distinction between speech and communication at an earlier stage. Our article is not specifically about language impairment. As a result of introducing the ICF framework earlier here, we can remove text from later in the introduction where we introduced ICF levels. See tracked changes.

Comment: how closely related difficulties" IN "speech"?

Response: This sentence has

Comment: "iv) also how wider aspects of communication such as emotional conveyance may be affected."- This seems to come out of nowhere and should be justified earlier. Why did you choose to look at emotional conveyance specifically?

Response: We have deleted the paragraph in question and replaced its content with a clearer and better placed statement of our research questions at the end of the introduction. We consider that this has improved the flow of ideas in the introduction and means that the emotional conveyance aspect emerges more naturally from our considerations about communication, making it clearer as to why we included emotional conveyance. This change of structure places greater emphasis on the sentence "Moreover, the ability to communicate emotions effectively is important in everyday life and studies have shown that reduced pitch variation and facial expression can cause negative evaluations of the personality of people with PD.", thereby strengthening the presentation of our rationale for this aspect of the study

Comment: "and also be an important predictor" – change: and is an important predictor

Response: We made this change when addressing grammatical issues in the manuscript. See tracked changes.

Comment: Sentence beginning: "A systematic review..." is grammatically flawed. Revise.

Response: We have revised the grammar of this sentence. See tracked changes.

Comment: "None of the studies had used... a validated outcome measure that assessed either communicative effectiveness of communicative participation as a concept, rather than specific sub aspects..." Why is this important? Needs further justification.

Response: We have added text at this point to clarify the importance of these methodological shortcomings in previous research. This added text reads "Therefore, these studies may have failed to detect mild cognitive impairment short of dementia and also to accurately capture the concept of functional communication, resulting in potential inaccurate measurement of both independent and dependent variables."

Comment: Does ICF need to be defined?

Response: It is defined and a reference provided at its first mention in the text, which is now earlier than in the original submission. See tracked changes.

Comment: Remove: "...the practicalities of such a large sample size probably explain why"...

Response: We have made the requested change. See tracked changes.

Comment: "In the only study to assess normal listeners' ability to identify specific emotions in the speech of people with PD..." Please also discuss here Pell, Cheang & Leonard, 2006.

Response: We have added a reference to the study by Pell et al, 2006, which we agree is relevant. We have re-phrased and added to this section of the text accordingly. See tracked changes. We have added the following reference:

- Pell MD, Cheang HS, Leonard CL. The impact of Parkinson's disease on vocal-prosodic communication from the perspective of listeners. *Brain Lang* 2006; 97: 123-34.

Comment: At the end of the introduction please provide an in-depth overview of your exact research questions and hypotheses.

Response: See response to "The manuscript would benefit considerably from an explicit description of the research questions addressed and hypotheses tested." above.

Comment: Method:

First section: please say more here about what the cross sectional design was actually assessing.

Response: The first part of the methods section in the BMJ Open style is about the study design. We have added the sub-heading "Design" to make this clearer. However, we have also re-phrased the first sentence to say what the study aim was, as you request, but this is not the main focus of this part of the BMJ Open article structure. We have added the text "In order to assess associations along the potential pathway to functional communication difficulties in PD,"

Comment: Participants: How many PD participants and CPs were recruited??

Response: According to community standards for medical journals, information about the participants actually recruited, rather than the procedures by which they are recruited, should be in the results section rather than the methods section. At the start of the 'Participants' part of the results section, we have added the following text "Forty five people with PD and 29 CPs were recruited"

Comment: Critical - How was the sample of 20 PD and 20 CP chosen from the larger sample?

Response: We have added the following text as explanation: "In order to generate our purposive sample, firstly, any samples that suffered from technical failure, other issues such as road noise and non-compliance with the task instructions were included. Then, selection sought to achieve a balanced profile of demographic and clinical features among people with PD and maximise comparability of demographics between the PD and CP groups, within the bounds of what was available in our sample".

Comment: "Since CPIB was our primary measure..." Why? Justify earlier.

Response: We have added text at the end of the introduction regarding functional communication being our primary outcome. At this point in the methods section, to justify why CPIB was selected as our primary measure of functional communication, we have added: "CPIB was chosen as our primary measure of functional communication since it specifically assesses ICF participation level difficulties that have been shown to be most important to people with PD, and also has been thoroughly developed using item-response theory methods and subsequent validated in PD in the United States and New Zealand, which are English-speaking countries." We have adjusted the structure of the surrounding sentence to fit the new content.

Comment: What were the 3 sentences from Miller and the one additional sentence read by the participants?

Response: We have now rephrased this section to include these 4 sentences. We have now requested and received copyright clearance to include Nick Miller's sentences in our manuscript.

Comment: Was the reliability check conducted on 10% of the data from every participant, or 100% of the data from 10% of the participants? If the latter, this could be problematic since the acoustics of people with Parkinson's will vary considerably from person to person. Discuss.

Response: It was 10% of the entire dataset selected at random. We agree that it would not have been ideal to re-check 100% of the data from 10% of participants, and therefore chose not to take that approach. We considered a 10% random selection of all acoustic data points across the full range of participants most appropriate. We have re-phrased to clarify how stimuli were selected for the reliability check. The additional text reads "of a randomly selected 10% sample of acoustic data points"

Comment: "In the intelligibility task, all stimuli were presented audiovisually, while in the emotional conveyance tasks, half were presented audiovisually and half in audio only." Why? Justify this decision.

Response: We have added text to clarify the rationale for this decision. The additional text reads "The rationale for including an audio-only condition in the emotional conveyance assessment was to test

the preliminary finding by Miller et al that that listeners were less likely to correctly identify the intended emotion in the speech of people with PD when auditory and visual information were both available. In contrast, for intelligibility assessment, we wanted to replicate the most common real-life listening conditions through presenting audiovisual information.”

Comment: On what basis was the predicted effect size $r=.05$ chosen for use in the power analysis? Was this based on the result of previous studies? Please cite them here.

Response: Ours is an exploratory study of an area (functional communication in Parkinson’s disease) that has received relatively limited research attention, and studies have differed greatly in terms of their exposure and outcome measurement, as well as issues of design. Therefore, unlike in an RCT, it is not possible to highlight one particular study as the basis of the expected effect size. Rather, it is based on several factors. An important factor is the extant body of literature. Preliminary systematic literature searches, which were the inspiration for what became our systematic review (cited as reference 23), indicated that the relationship was of a moderate nature in the existing literature. However, in appraising the body of literature, we noted that no study had used combinations of exposure and outcome measures that could be taken alone in informing a power calculation. Another important factor in our decision was our combined wider theoretical, scientific and clinical knowledge and experience about communication in neurological disorders such as Parkinson’s disease. This corroborated our understanding of the limited existing literature, and informed us that functional communication in Parkinson’s disease is an area upon which many factors bring to bear, among which cognitive status is an important factor, but far from the only one. Therefore, our understanding of the limited existing body of literature on our specific research question, combined with our wider understanding and experience of the topic, informed us that the expected effect size for our primary relationship between cognitive status and functional communication in Parkinson’s disease would be of a moderate nature. We then brought this body of information to our study steering committee, and in particular our study’s senior statistician (author ABC, who is a Senior Lecturer (Associate Professor) in Medical Statistics) for consideration. Author ABC then advised us on the basis of his expertise and experience that the most appropriate number to use in the power calculation would be $r=0.5$. We have added text to this situation to briefly summarise our rationale. We consider that using a power calculation, albeit having to make a judgement over the exact effect size to enter, is a strength of our study compared to previous studies in this field. We report in our systematic review (as cited in the manuscript) that the lack of any reported power calculation or sample size rationale was a common methodological limitation across included studies. The additional text reads “The effect size to use for the power calculation was determined by senior statistician A.B.C. informed by i) preliminary systematic literature searches by the research team that later became our systematic review and ii) the research team’s combined wider theoretical, scientific and clinical knowledge and expertise about communication in neurological conditions such as PD, which both informed us to expect a moderate relationship between cognitive status and functional communication in PD.”

Comment: Please report the results of the test-retest reliability and convergent validity of CPIB in the method not the results.

Response: Community standards for medical journal articles ask for all results, that is to say what is not known before the study commences, to be put in the results section rather than the methods section. Therefore, we have retained the results of test-retest reliability and convergent validity in the results section.

Comment: How was motor impairment assessed? UPDRS, Hoehn and Yahr etc.

Response: We have added text to clarify that Levodopa Equivalent Daily Dose (LEDD) was used as a

proxy measure of non-speech-specific motor symptom severity. We did not have access to clinician-administered measures such as UPDRS and Hoehn and Yahr. Our steering committee advised that LEDD was an appropriate proxy measure. The additional text reads “LEDD served as a proxy measure of non-speech-specific PD motor symptom severity.”

Comment: Were transcribers allowed to listen to the sentence only once or more than that? If once, were the sentences of the two groups matched in length?

Response: We have added text to state “In all listener assessment tasks, assessors could only listen to each sentence once and sentences from people with PD and CPs were matched for length.”

Results:

Comment: Why is smoking status described? Please describe earlier why this information was collected and what it tells us.

Response: We agreed that this information does not add to the manuscript and have therefore removed it, since this manuscript is too long to carry unnecessary information

Comment: Please include the exact test statistic, p value, effect size and 95% confidence interval for all between group comparisons reported in the results.

Response: Here we have a situation in which the two Reviewers disagree. We note that the Editor made no comment on the presentation of results. Moreover, as stated in the manuscript, there are no standard reporting guidelines for studies such as ours. Reviewer One (Dr Fogg-Rogers) said “The results and statistics are clearly presented with well defined key findings in the discussion”. We have thoroughly considered the perspectives of both Reviewers, and decided to retain our current results presentation in line with the comments of Reviewer One. The analysis plan, including the way in which our results are reported, was devised by our senior statistician author A.B.C. following detailed and thorough discussion with the investigator team and the steering committee. Subsequent to your comment, the corresponding author M.S.B, who also has statistical expertise, has thoroughly considered and reviewed all decisions made by our senior statistician author A.B.C. with regard to how the results are reported for our primary and secondary outcomes respectively. Furthermore, further statisticians who are not involved in the current study have also provided their perspective on this matter to the corresponding author. In particular, the statisticians agreed with us that the use of 95% confidence intervals would not be appropriate in our study. The rationale was that this is a wide-ranging exploratory study to identify effects for investigation in future more targeted confirmatory studies. Confidence intervals will be valuable to use in these confirmatory studies to indicate the precision of confirmed estimates. However, to use confidence intervals in a wide-ranging exploratory study would be unhelpful and provide spuriously precise estimates of certainty, which are not the purpose of this study. Following these considerations, there was unanimous agreement in favour of retaining the way results were presented in the original manuscript, following the comment of Reviewer One.

Comment: Was there an association between % words correctly transcribed, and emotional conveyance?

Response: In response to your previous comment, we have clarified our research questions. The association between intelligibility and emotional conveyance does not form part of our study, but may be an area for future research if others consider this a priority area. Due to the wide range of research questions in our study and the length of our manuscript, it is not possible to add further off-protocol research questions at this time.

Comment: Discussion:

In the emotional conveyance section please also discuss your findings with relation to Pell, Cheang and Leonard, 2006. Why would the happy emotion be particularly affected?

Many of your findings were more strongly associated with gender than with PD. Please discuss.

Response: In the first sentence of the emotional conveyance section of the discussion, we have added “and Pell et al” directly after our reference to Miller. Regarding why happy emotion was particularly affected, our study cannot draw conclusion on this and we have added clarification to this regard: “, although our study cannot confirm the mechanisms which might be causing this effect”. With regard to gender, we add text to emphasise the need for further research to clarify the gender-specific effects of Parkinson’s on communication, since this has received limited research attention: “In addition, it is important to conduct further research into the gender-specific aspects of communication difficulties in PD, which have received limited research attention”

Comment: “Intelligibility did not account for a large proportion of variance in functional outcomes” – was it not the case that intelligibility predicted both communicative participation and effectiveness?

Response: Intelligibility in conversational sentences was a statistically significant predictor of functional communication outcomes. Intelligibility in read sentences was not a statistically significant predictor of functional communication outcomes. Although statistically significant, the proportion of variance in functional communication outcomes explained by conversational sentence is relatively small. Therefore, we are satisfied that this sentence is an accurate reflection of the results of our study.

Comment: As far as I can tell, level of motor disability was not assessed. Since greater cognitive impairment is likely to be strongly correlated with greater motor impairment (as both decline in parallel), can you truly say that cognitive factors are more important than motor factors (i.e. could there be a confound here)? Despite the fact that there were very few differences between the two groups on your acoustic measures, it was still the case that patients were significantly less intelligible. This indicates that motor speech impairments still exerted a strong general effect, even though this was not captured in the fine grained analysis of different types of speech acoustics. Could it have made for a stronger study design to assess motor disability more directly to rule this out? Please discuss.

Response: We have added the following text to the limitations section of the discussion: “Fifthly, we were unable to measure motor disability directly. However, we offered LEDD as a proxy measure of motor disability to models assessing functional outcomes in order to minimise confounding by motor disability.” We did not want to side-track the paper here into a discussion of the relationship between motor and cognitive impairment in Parkinson’s disease, which is not the focus of our paper. We have, however, added text to clarify “Moreover, some studies have shown cognitive impairment can be common in people with PD who are early on the motor decline pathway”, which suggests that cognitive and motor disability may not always parallel each other. We have now additionally cited

- Yarnall AJ, Breen DP, Duncan GW, et al. Characterizing mild cognitive impairment in incident Parkinson disease. *Neurol* 2014; 82: 308-16.
- Pedersen KF, Larsen JP, Tysnes OB, et al. Prognosis of mild cognitive impairment in early Parkinson disease: the Norwegian ParkWest study. *JAMA Neurol* 2013; 70: 580-6.

VERSION 2 – REVIEW

REVIEWER	Stacey Humphries Postdoctoral Scholar, University of Pennsylvania, USA
REVIEW RETURNED	03-Jan-2017

GENERAL COMMENTS	<p>I thank the authors for their time and care in responding to my comments. The introduction has been considerably strengthened, particularly the rationale and aims/research questions. I have some remaining comments, most of which are aimed at increasing transparency.</p> <p>In the research questions section of the introduction, change “we decided to conduct” to “we conducted”.</p> <p>With regard to how the smaller sample of 20 PD and 20 CP were chosen from the larger sample, did you mean that technical failure samples were excluded, not included? You state that these participants were selected with the goal of achieving a balanced profile of clinical and demographic features. Can you provide more information about this? Does a balanced profile mean that you sought to maximize variability amongst the PD group, or minimize it? Which demographic and clinical features did you take into account when choosing who to include?</p> <p>10% of the data points across the entire data set were reanalysed as part of the reliability check. How many participants had at least one of their data points included in this 10%?</p> <p>Thank you for clarifying that LEDD was used as a proxy measure of motor symptom severity. Given the vast heterogeneity in patients in terms of how well they respond to medication, and whether or not they opt to go un-medicated in the early stages, I think it would be worth briefly mentioning the limitations of this measure. For example, Yoritaka et al., 2013, found very similar LEDD in patients at all 6 Hoehn and Yahr stages. Future research may want to follow-up your work by investigating some of the specific motor subscales from the UPDRS relating to voice and hand movements, for example, and how they relate to functional communication.</p> <p>With regard to the reporting of statistics, I take your point about confidence intervals. However, it is still my strong view that exact test statistics, p values and effect sizes/regression coefficients should be reported. Given the current replication crisis in medicine and psychology, open and transparent reporting of statistics is of great importance. Also, since you point out that your study is exploratory, one would hope that others will follow up on your work in the future. Transparent reporting of statistics will be useful in that regard, because mean differences are difficult to interpret in isolation. Furthermore, just knowing which tests produced a significant result is not sufficient, since there were so many tests conducted it is possible that some significant findings may be the result of a type 1 error. That said, I understand that the journal does not have specific requirements with regard to the reporting of statistics so I will defer to the editor on this.</p> <p>Asako Yoritaka, Yasushi Shimo, Masashi Takanashi, Jiro Fukae, Taku Hatano, Toshiki Nakahara, Nobukazu Miyamoto, Takao Urabe, Hideo Mori, Nobutaka Hattori, Motor and non-motor symptoms of</p>
-------------------------	---

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2

Reviewer Name: Stacey Humphries

Institution and Country: Postdoctoral Scholar, University of Pennsylvania, USA Competing Interests: None declared

Comment: I thank the authors for their time and care in responding to my comments. The introduction has been considerably strengthened, particularly the rationale and aims/research questions. I have some remaining comments, most of which are aimed at increasing transparency.

Response: We thank the reviewer for this comment. We have sought to clarify these remaining comments below

Comment: In the research questions section of the introduction, change “we decided to conduct” to “we conducted”.

Response: We have made the suggested change on the two occasions this phrase appeared in the research questions section of the introduction

Comment: With regard to how the smaller sample of 20 PD and 20 CP were chosen from the larger sample, did you mean that technical failure samples were excluded, not included? You state that these participants were selected with the goal of achieving a balanced profile of clinical and demographic features. Can you provide more information about this? Does a balanced profile mean that you sought to maximize variability amongst the PD group, or minimize it? Which demographic and clinical features did you take into account when choosing who to include?

Response: We have clarified the requested matters. Technical failure samples were excluded. We have corrected this in the text. We have further clarified our purposive sampling. The following text has been added: “Only people with PD who provided a CP were considered. Age, gender, accent and perceived severity of speech disorder were also considered in selection. In particular, it was important to ensure generalisability of the PD sample. ” So, we did not seek to either maximise or minimise variation but generate a sample that is as well matched as we could achieve to the wider PD population

Comment: 10% of the data points across the entire data set were reanalysed as part of the reliability check. How many participants had at least one of their data points included in this 10%?

Response: We have provided the requested additional information. It was a 10% random check of all data points in the sample. Our phonetician Z.R.B. has confirmed that this included data points from 10 participants, that is to say 25% of the sample size used for phonetic analysis. We have added text to clarify this: “drawn from 10 different participants (25% of the phonetic analysis sample size)”

Comment: Thank you for clarifying that LEDD was used as a proxy measure of motor symptom severity. Given the vast heterogeneity in patients in terms of how well they respond to medication, and whether or not they opt to go un-medicated in the early stages, I think it would be worth briefly mentioning the limitations of this measure. For example, Yoritaka et al., 2013, found very similar

LEDD in patients at all 6 Hoehn and Yahr stages. Future research may want to follow-up your work by investigating some of the specific motor subscales from the UPDRS relating to voice and hand movements, for example, and how they relate to functional communication.

Asako Yoritaka, Yasushi Shimo, Masashi Takanashi, Jiro Fukae, Taku Hatano, Toshiki Nakahara, Nobukazu Miyamoto, Takao Urabe, Hideo Mori, Nobutaka Hattori, Motor and non-motor symptoms of 1453 patients with Parkinson's disease: Prevalence and risks, *Parkinsonism & Related Disorders*, Volume 19, Issue 8, August 2013, Pages 725-731

Response: We have added additional text as requested to say: "However, LEDD has limitations as a proxy measure of motor status. For example, one study found no significant association between LEDD and Hoehn and Yahr staging. Therefore, future studies should consider assessing how scores from explicit motor assessments, such as the Universal Parkinson's Disease Rating Scale (UPDRS) predict functional communication outcomes. ". We have also cited Yoritaka et al as requested, and cited a paper for UPDRS.

Comment: With regard to the reporting of statistics, I take your point about confidence intervals. However, it is still my strong view that exact test statistics, p values and effect sizes/regression coefficients should be reported. Given the current replication crisis in medicine and psychology, open and transparent reporting of statistics is of great importance. Also, since you point out that your study is exploratory, one would hope that others will follow up on your work in the future. Transparent reporting of statistics will be useful in that regard, because mean differences are difficult to interpret in isolation. Furthermore, just knowing which tests produced a significant result is not sufficient, since there were so many tests conducted it is possible that some significant findings may be the result of a type 1 error. That said, I understand that the journal does not have specific requirements with regard to the reporting of statistics so I will defer to the editor on this.

Response: We thank Reviewer 2 for her interest in our statistical analysis and the Editor for his guidance. We have gone through the results section of the manuscript and the results section of the abstract and added exact p-values instead of 'less than' p-values for all instances in which these occurred. The only exception is when the p-value says $p < 0.001$. There is a strong community convention in medical statistics to report these values as $p < 0.001$ rather than $p = 0.00$ or $p = 0.000$. With regard to the comment about effect measures, we went through the manuscript carefully to check that all were reported as exact values. We replaced any mean difference scores that were presented as integers with values showing decimal place. We also took the opportunity to ensure that the results were presented in exactly the same way in the main manuscript and in the abstract.

As for the Supplementary tables, the Editor did not mention these (instead referring to the results section) and they serve a different purpose. They are intended to provide an at-a-glance view of a wide range of information that would not fit into the text of the manuscript. Presentation and readability are key here, and we consider it important to retain clarity of presentation. It is standard practice in such tables to indicate levels of statistical significance using asterisks (* = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$) as we have done. Such tables do not afford the space to add exact p-values.

We thank the Reviewer and the Editor for their input, and hope that our clarification and additional of further details where appropriate is acceptable to the Editor.

VERSION 3 – REVIEW

REVIEWER	Stacey Humphries University of Pennsylvania, USA
REVIEW RETURNED	17-Feb-2017

GENERAL COMMENTS	I am satisfied with the authors' responses to my previous review.
-------------------------	---