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)	Long term use benefits of personal frequency-modulated systems for speech in noise perception in stroke patients with auditory processing deficits: a non randomised controlled trial study
AUTHORS	Koohi, Nehzat; Vickers, Deborah; Warren, Jason; Werring, David; Bamiou, Doris-Eva

VERSION 1 - REVIEW

REVIEWER	Eliane Schochat Universidade de São Paulo - Brazil
REVIEW RETURNED	29-Jun-2016

GENERAL COMMENTS	The article is very well written and the subject of the study very interesting and update. However I have some comments: - First, I think that the authors should explain more in depth the use of the FM system specially at the social situations. - Second I think that the statistical description is not clear. It should be more well described. Finally, I think that the sentence on page 14, line 8/10 ("This updating of predictions at higher level is heavily dependent on salience of stimulus and context, i.e. attention") should be explained more in depth because I feel that the conclusion of the sentence "i.e. attention" is far from the beginning of the sentence. And less important, on page 2 line 33/34 when is written "speech in noise symptoms" I would suggest "speech in noise problems".

REVIEWER	Dr. Ajith Kumar U All India Institute of Speech and Hearing Mysore
REVIEW RETURNED	05-Jul-2016

GENERAL COMMENTS	This study looks at long-term benefits of use of FM devices in stroke patients with auditory processing deficits. My major concerns with the study are: 1) PAT provides average thresholds. But SPIN heavily depends on high frequency hearing status. So it is important to provide and control this affect. 2) Looking at the ages of the participants it appears that some of the participants may have had age related high frequency hearing loss and this could have affected the results. 3) It is also well known that high frequency hearing loss adversely affects auditory processing test results. Therefore, providing complete audiological details of the participants is essential to interpret the results of the study. There fore results of theses study

	could be interpreted only if participants did not have high frequency hearing loss. 4) I am also not sure of statistical treatment. As sample size is too low and I doubt data has met the assumptions of GLM, such as normality etc. According to me data needs to re-analyzed with non-parametric tests for interpretation. 5) As sample size is too low, individual data/improvement graphs are essential to draw any conclusions. Because of these reasons, though study has some merit in treatment of stroke patients, I believe that data provided and statistical treatment done are insufficient and inappropriate to draw any.

REVIEWER	Gordon S. Doig University of Sydney, Australia
REVIEW RETURNED	09-Aug-2016

GENERAL COMMENTS	Review of manuscript Bmjopen-2016-01300 titled “Proof of concept study on long term use benefits of personal frequency-modulated systems for speech in noise perception in stroke patients with auditory processing deficits.” Please address the following issues with changes or additions to your manuscript: 1. Title: Please alter your title to reflect your study design using three concise words at the end of the current title. 2. Study design: This is NOT a case-control study. This is NOT an observational study, because you ASKED patients if they would be willing to try the study intervention. Your study is most likely best described as a ‘non-randomised controlled trial’. Please use this description in your Title, Abstract and Methods. In the second paragraph of your Discussion, please warn the reader of the severe types of biases that can arise due to non-random allocation. 3. Because you did not randomise patients to receive the intervention, you have not established that patients would be willing to accept randomisation to your study groups. This is a very important concept that must be addressed in your Discussion. 4. Change from baseline may appear to be occurring in one group if baseline levels are higher in that one group, resulting in regression towards the mean. Please establish that baseline levels in both groups are similar and outcome levels (not change scores) are similar. 5. Because you did not randomise, you MUST provide a covariate adjusted analysis. 6. Page 5. Is this a ‘proof of concept’ pilot study or a very small interventional study? If it is proof of concept, your hypothesis should be about addressing study dynamics (completion rate etc), not about treatment effects. 7. Please write in complete sentences, not –point form.

	8. Whenever you report a %, please report numerator / denominator immediately afterwards. 9. Please do not refer to your study participants as ‘cases’ or ‘controls’. They are ‘intervention subjects’ or ‘standard care’ subjects. 10. Page 11. For your primary outcome, please report baseline values and week 10 values, with measures of variance and simple p-values for each time point. I cannot interpret the validity of a GLM based on 9 participants with your current level of reporting. Your model may be too complex for your sparse data. 11. Table 1 and Figures. It looks as if one outlier is driving your results. 12. This is an interventional study. Did you list it with a Registry before you began? 13. Your Discussion must start out by addressing issues of feasibility, if this is a ‘proof of concept’ study. Very early in your Discussion, you must warn your readers about the limitations of your extreme small size AND failure to randomise. You CANNOT just state that ‘we have small size and failed to randomise’. Please cite and interpret key methodological references that communicate the appropriate caution that should be used to interpret any information arising from this VERY small trial. I would recommend you remove ALL comments on treatment effects, physiology and conclusions. And just focus on the need to validate your preliminary findings in a large well conducted trial.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1.
“the authors should explain more in depth the use of the FM system specially at the social situations”
We added this paragraph: “Patients were asked to use the FM at home with family members and with multiple media devices such as music players, radio, the television, and the computer. They were also asked to use the FM in social situations such as family outings (with their significant others wearing the microphone), restaurants, public houses, as well as small group meetings at work etc.”
“ the statistical description is not clear. It should be more well described” this section has been entirely rewritten to address comments by all reviewers
“the sentence on page 14, line 8/10 should be explained more in depth”
This was rephrased to “This updating of predictions at higher level is heavily dependent on attention mechanisms influenced by the degree of salience of the stimulus and the listening context “

"speech in noise symptoms" I would suggest "speech in noise problems". This has been rephrased as “reported difficulties understanding speech in background noise”
Reviewer: 2
“PAT provides average thresholds. But SPIN heavily depends on high frequency hearing status. So it is important to provide and control this affect Looking at the ages of the participants it appears that some of the participants may have had age related high frequency hearing loss and this could have affected the results. It is also well known that high frequency hearing loss adversely affects auditory processing test results. Therefore, providing complete audiological details of the participants is essential to interpret the results of the study. There fore results of theses study could be interpreted only if participants did not have high frequency hearing loss. ” We have provided the high frequency

PTA average and high frequency average for intervention and standard care subjects, which were not significantly different. Auditory processing tests used for this study included the gaps in noise test- which is conducted at 50 dB sensation level at 1 kHz (ie a threshold that was normal for all cases and controls) and the Queen's Square Tests of Auditory Cognition (QSTAC) auditory processing battery, which is relatively resilient to peripheral hearing loss (see Goll et al., 2010). We have not provided these results in detail in this paper since

- Summary results of these patients as part of a bigger stroke cohort vs controls are given in a paper under review by JAAA
- Individual and group test results of these patients as part of a bigger stroke cohort vs their stroke lesion are discussed in a further paper under preparation.

"I am also not sure of statistical treatment. As sample size is too low and I doubt data has met the assumptions of GLM, such as normality etc. According to me data needs to re-analyzed with non-parametric tests for interpretation" we have used bootstrapping techniques to address the small sample size and acknowledged these limitations in the discussion.

" As sample size is too low, individual data/improvement graphs are essential to draw any conclusions. " revised figure 1 provides individual data at visit 1 and visit 2 in case and control subjects, in aided and unaided conditions. We have also improved the statistical methods to address these concerns.

Reviewer: 3

"Your study is most likely best described as a 'non-randomised controlled trial'. Please use this description in your Title, Abstract and Methods"

This has been done

"In the second paragraph of your Discussion, please warn the reader of the severe types of biases that can arise due to non-random allocation. .. Because you did not randomise patients to receive the intervention, you have not established that patients would be willing to accept randomisation to your study groups. This is a very important concept that must be addressed in your Discussion. "

We have added this into discussion "The study has small case numbers, and findings ought to be interpreted with caution. Low study numbers affect precision of measurements, and even a nominally statistically significant finding may not reflect a true effect. 23 Furthermore, groups were not randomly allocated, thus being prone to selection bias, while willingness for randomization and the factors influencing patient preference and thus group allocation 24 were not determined. An additional limitation was that we could not investigate the effect of extent of brain lesion on results. However, the study purpose was to determine if there is potential value in running a large scale intervention trial."

"Change from baseline may appear to be occurring in one group if baseline levels are higher in that one group, resulting in regression towards the mean. Please establish that baseline levels in both groups are similar and outcome levels (not change scores) are similar" this has been done

"Because you did not randomise, you MUST provide a covariate adjusted analysis" we conducted bootstrapped one-way analysis of covariance in the revised manuscript.

"Page 5. Is this a 'proof of concept' pilot study or a very small interventional study? "This was a small intervention study. We also collected additional information being conducted about feasibility aspects.

"Please write in complete sentences, not –point form" This has been done with the exception of test and test parameters outline.

"Whenever you report a %, please report numerator / denominator immediately afterwards" This has been done for % pertaining to this present paper

"Please do not refer to your study participants as 'cases' or 'controls'. They are 'intervention subjects' or 'standard care' subjects. " This has been done

" For your primary outcome, please report baseline values and week 10 values, with measures of variance and simple p-values for each time point." This has been done in table 2

" I cannot interpret the validity of a GLM based on 9 participants with your current level of reporting. Your model may be too complex for your sparse data." We have reanalysed data with a bootstrapping

ANCOVA
“Table 1 and Figures. It looks as if one outlier is driving your results” please see new figure 1 which provides individual data
“This is an interventional study. Did you list it with a Registry before you began? “ this study was registered after the end of the study with ClinicalTrials.gov.
“Your Discussion must start out by addressing issues of feasibility, if this is a ‘proof of concept’ study”. We report this as a small intervention study.
“Very early in your Discussion, you must warn your readers about the limitations of your extreme small size AND failure to randomise. You CANNOT just state that ‘we have small size and failed to randomise’. Please cite and interpret key methodological references that communicate the appropriate caution that should be used to interpret any information arising from this VERY small trial”
As previously stated, we added this referenced paragraph early in the discussion:
““The study has small case numbers, and findings ought to be interpreted with caution. Low study numbers affect precision of measurements, and even a nominally statistically significant finding may not reflect a true effect. 23 Furthermore, groups were not randomly allocated, thus being prone to selection bias, while willingness for randomization and the factors influencing patient preference and thus group allocation 24 were not determined. An additional limitation was that we could not investigate the effect of extent of brain lesion on results.”
“ I would recommend you remove ALL comments on treatment effects, physiology and conclusions. And just focus on the need to validate your preliminary findings in a large well conducted trial.” We have limited discussion of results physiology, however we do feel that some tentative explanation needs to be proposed. We eliminated conclusions other than the need for a rigorous RCT.

VERSION 2 – REVIEW

REVIEWER	Eliane Schochat Universidade de São Paulo - Medical School
REVIEW RETURNED	04-Sep-2016

GENERAL COMMENTS	- The authors have done a great job with the manuscript reviews and I feel that my comments have been fully accommodated in the latest revision. I am pleased to recommend this manuscript for publication. The manuscript is enhanced by the more detailed mechanistic consideration of use of the FM system and the statistical analizes.

REVIEWER	Gordon S. Doig University of Sydney
REVIEW RETURNED	16-Sep-2016

GENERAL COMMENTS	My original comments were numbered sequentially, but the author responses are not numbered and do not appear to address my comments sequentially. In your response, please provide each of my numbered comments, numbered in sequence, with each comment reported in full (not edited). Please provide a complete response to each comment immediately after each comment. Sorry, but I cannot afford the time to search through your out of sequence responses to ensure each of my comments have been appropriately addressed.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 3

1. “Title: Please alter your title to reflect your study design using three concise words at the end of the current title.”

This has been changed to (see final segment after the hyphen) Long term use benefits of personal frequency-modulated systems for speech in noise perception in stroke patients with auditory processing deficits- a non randomised controlled trial study

2.” Study design: This is NOT a case-control study. This is NOT an observational study, because you ASKED patients if they would be willing to try the study intervention. Your study is most likely best described as a ‘non-randomised controlled trial’. Please use this description in your Title, Abstract and Methods”

This has been done, see title above, abstract and tracked changes in manuscript in section methods, study design. Subjects were classified into intervention subjects (who received the FM) and standard care subjects (who did not receive the FM).

“ In the second paragraph of your Discussion, please warn the reader of the severe types of biases that can arise due to non-random allocation.”

This was done in the first paragraph’s second part:

“The study has small case numbers, and findings ought to be interpreted with caution. Low study numbers affect precision of measurements, and even a nominally statistically significant finding may not reflect a true effect. 23 Furthermore, groups were not randomly allocated, thus being prone to selection bias, while willingness for randomization and the factors influencing patient preference and thus group allocation 24 were not determined. An additional limitation was that we could not investigate the effect of extent of brain lesion on results. However, the study purpose was to determine if there is potential value in running a large scale intervention trial.”

3. “Because you did not randomise patients to receive the intervention, you have not established that patients would be willing to accept randomisation to your study groups. This is a very important concept that must be addressed in your Discussion.”

We added this sentence in the discussion: “ … and findings ought to be interpreted with caution. …Furthermore, groups were not randomly allocated, thus being prone to selection bias, while willingness for randomization and the factors influencing patient preference and thus group allocation 24 were not determined. “

4.” Change from baseline may appear to be occurring in one group if baseline levels are higher in that one group, resulting in regression towards the mean. Please establish that baseline levels in both groups are similar and outcome levels (not change scores) are similar.”

this has been addressed: please see table 2, also see the following sentence in the new statistical analysis section “A series of bootstrapped ANCOVAs were calculated to examine the differences between the intervention and comparison groups on visit 2 (post intervention) scores, while controlling for visit 1 (baseline) scores of the same measure.”

5. “Because you did not randomise, you MUST provide a covariate adjusted analysis.”

We conducted bootstrapped one-way analysis of covariance in the revised manuscript. Pl see statistical analysis section.

6. “Page 5. Is this a ‘proof of concept’ pilot study or a very small interventional study? If it is proof of concept, your hypothesis should be about addressing study dynamics (completion rate etc), not about treatment effects.”

This was a small intervention study. We also collected additional information being conducted about feasibility aspects. We were interested in retention and adherence and we provided the following sentence in results: “There were no drop outs to the study, and all recruited subjects came back for retest 10 weeks later. All 4 intervention subjects complied with the daily use of the FMs.”

7.” Please write in complete sentences, not –point form.”

This has been done with the exception of test and test parameters outline.

8. “Whenever you report a %, please report numerator / denominator immediately afterwards.”

This has been done for % pertaining to this present paper, with the exception of a percentage value quoted in reference 7.

9. “Please do not refer to your study participants as ‘cases’ or ‘controls’. They are ‘intervention subjects’ or ‘standard care’ subjects.”

This has been changed throughout the manuscript

10. “Page 11. For your primary outcome, please report baseline values and week 10 values, with measures of variance and simple p-values for each time point.”

This has been done in table 2

“ I cannot interpret the validity of a GLM based on 9 participants with your current level of reporting. Your model may be too complex for your sparse data.”

We have reanalysed data with a bootstrapping ANCOVA

11. “Table 1 and Figures. It looks as if one outlier is driving your results.”

please see new figure 1 which provides individual data

12. “This is an interventional study. Did you list it with a Registry before you began?”

this study was registered after the end of the study with ClinicalTrials.gov. - NCT02889107

13. “Your Discussion must start out by addressing issues of feasibility, if this is a ‘proof of concept’ study.”

We report this as a small intervention study. The most important feasibility aspect was whether the study participants would be willing to use the FM and the discussion starts with the sentence: “Forty four percent of eligible stroke subjects for the FM intervention (i.e. with preserved peripheral hearing but with speech in noise reported difficulties and test deficits) were willing to use the FM systems daily.”

“Very early in your Discussion, you must warn your readers about the limitations of your extreme small size AND failure to randomise. You CANNOT just state that ‘we have small size and failed to randomise’. Please cite and interpret key methodological references that communicate the appropriate caution that should be used to interpret any information arising from this VERY small trial. “

We added this referenced paragraph early in the discussion: “The study has small case numbers, and findings ought to be interpreted with caution. Low study numbers affect precision of measurements, and even a nominally statistically significant finding may not reflect a true effect. 23 Furthermore, groups were not randomly allocated, thus being prone to selection bias, while willingness for randomization and the factors influencing patient preference and thus group allocation 24 were not determined. An additional limitation was that we could not investigate the effect of extent of brain lesion on results.”

“I would recommend you remove ALL comments on treatment effects, physiology and conclusions. And just focus on the need to validate your preliminary findings in a large well conducted trial.”

We have limited discussion of results physiology, however we do feel that some tentative explanation needs to be proposed. We are applying for funding to assess the mechanism for the observed FM unaided benefits. We eliminated conclusions other than the need for a rigorous RCT.

VERSION 3 – REVIEW

REVIEWER	Gordon S. Doig University of Sydney
REVIEW RETURNED	02-Nov-2016

GENERAL COMMENTS	Review of bmjopen-2016-013003.R1 titled “Long term use benefits of personal frequency-modulated systems for speech in noise perception in stroke patients with auditory processing deficits- a non randomised controlled trial study.” 1. The decision to bootstrap is complex, and I am not certain it is most appropriate for an extremely small dataset. Please provide methodological references to support the use of bootstrapped ANOVAs and ANCOVAs in your small study. Please provide more details about your bootstrapping process (number of interations, adjustment for bias) and please report the detailed results of a simple ANOVA and ANCOVA in Table format (to include F-statistics and p-values from F-tests, not p-values from post hoc contrasts). 10. “Page 11. For your primary outcome, please report baseline values and week 10 values, with measures of variance and simple p-values for each time point.” Please report simple p-values, not bootstrapped p-values. These 'simple' direct tests represent univariate analyses. Your covariate adjusted analysis represents your most important results. The direct comparison univariate analysis is simply 'supportive'.

VERSION 3 – AUTHOR RESPONSE

Thank you for your helpful comments. We have addressed these as follows:
1. "The decision to bootstrap is complex, and I am not certain it is most appropriate for an extremely small dataset. Please provide methodological references to support the use of bootstrapped ANOVAs and ANCOVAs in your small study. Please provide more details about your bootstrapping process

(number of interations, adjustment for bias) and please report the detailed results of a simple ANOVA and ANCOVA in Table format (to include F-statistics and p-values from F-tests, not p-values from post hoc contrasts)."

We provided this paragraph in the methods section, statistical analysis:
"...Bootstrapped one-way analysis of variance (ANOVA) and analysis of covariance (ANCOVA) were thus conducted (Fisher and Hall, 1991; Barber et al., 2000) to determine a statistically significant difference between the intervention and standard care groups on the speech reception thresholds in noise at the baseline time point. Bootstrapped confidence intervals (95% CIs, bias-corrected, accelerated with 1000 replications) were calculated. ANCOVA was performed to control for age and side of lesions, and to examine the differences between the intervention and standard care groups on post-intervention scores, while controlling for baseline scores of the same measure. The independent variable, study group, included two levels: intervention group and standard care group. The dependent variable was the SRT in noise scores and covariates were age and side of lesion."

We also provided a new table 2 with F statistics and p values for ANOVA and ANCOVA (see submission) and the following paragraph in results:
"The bootstrapped one-way analysis of variance (ANOVA) showed no statistically significant difference between the intervention and standard care groups on the speech reception thresholds in noise at visit 1 time point. There was no difference between groups with respect to the speech reception thresholds in noise when the babble was presented at either the left or the right side, in aided and unaided condition (Table 2). A series of bootstrapped univariate ANOVAs were calculated to examine the differences between the intervention and comparison groups on visit 2 (post intervention) scores. The bootstrapped ANOVA showed a statistically significant improvement in SRT in noise when the noise was coming from left and right for both aided and unaided conditions at visit 2 time point in the intervention group compared to those who received standard care. However, when a series of bootstrapped univariate analysis of covariance (ANCOVAs) were calculated to control for baseline outcomes, age and side of lesions, a statistically significant improvement in SRT in noise was observed only when the noise was coming from left for both aided and unaided conditions at visit 2 time point in the intervention group compared to those who received standard care (Table 2)."

10. “Page 11. For your primary outcome, please report baseline values and week 10 values, with measures of variance and simple p-values for each time point.” Please report simple p-values, not bootstrapped p-values. These 'simple' direct tests represent univariate analyses. Your covariate adjusted analysis represents your most important results. The direct comparison univariate analysis is simply 'supportive'.

We provided the simple direct test results in table 3 (see submission) and the following paragraph in results section:
"A series of Mann-Whitney U tests were performed in the different speaker testing positions, these were for baseline visit and 10-week visit to assess differences in the primary outcomes of SRT in noise between intervention and standard care groups. The results are summarised in Table 3."

VERSION 4 – REVIEW

REVIEWER	Gordon S. Doig University of Sydney, Australia.
REVIEW RETURNED	05-Dec-2016

GENERAL COMMENTS	Review of bmjopen-2016-013003.R1 titled “Long term use benefits of personal frequency-modulated systems for speech in noise perception in stroke patients with auditory processing deficits- a non randomised controlled trial study.”

	1. The decision to bootstrap is complex, and I am not certain it is most appropriate for an extremely small dataset. Please provide methodological references to support the use of bootstrapped ANOVAs and ANCOVAs in your small study. Please provide more details about your bootstrapping process (number of interations, adjustment for bias) and please report the detailed results of a simple ANOVA and ANCOVA in Table format (to include F-statistics and p-values from F-tests, not p-values from post hoc contrasts). In your response to this request, you refer me to Table 2. I cannot find F-statistics and p-values from the requested 'simple (not bootstrapped)' ANOVA and ANCOVA that I requested. Please provide these details in addition to your 'bootstrapped' results.

VERSION 4 – AUTHOR RESPONSE

1. The decision to bootstrap is complex, and I am not certain it is most appropriate for an extremely small dataset. Please provide methodological references to support the use of bootstrapped ANOVAs and ANCOVAs in your small study. Please provide more details about your bootstrapping process (number of interations, adjustment for bias) and please report the detailed results of a simple ANOVA and ANCOVA in Table format (to include F-statistics and p-values from F-tests, not p-values from post hoc contrasts).

In your response to this request, you refer me to Table 2. I cannot find F-statistics and p-values from the requested 'simple (not bootstrapped)' ANOVA and ANCOVA that I requested. Please provide these details in addition to your 'bootstrapped' results.

Authors' response: In our previous manuscript we used SPSS package for the data analysis, however SPSS is unable to calculate the bootstrapped F-statistic and bootstrapped p-values and only computes bootstrapped standard error and confidence intervals (the output for both simple and bootstrapped is same). After a consultation with a statistician, we decided to re-analyse our data in the STATA package in order to obtain both simple and bootstrapped F-statistics and p-values. Please refer to the results section (changes are highlighted in red text) and tables 2 and 3. Table 2 shows the simple ANOVA and ANCOVA with F-statistics and p-values. Table 3 shows the bootstrapped ANOVA and ANCOVA with standard error, z-statistics and p-values.

VERSION 5 – REVIEW

REVIEWER	Gordon S. Doig University of Sydney
REVIEW RETURNED	20-Jan-2017

GENERAL COMMENTS	The authors have more than adequately addressed my comments.