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)	Organising Support for Carers of Stroke Survivors (OSCARSS): a cluster-randomised controlled trial with economic evaluation
AUTHORS	Patchwood, Emma; Woodward-Nutt, Kate; Rhodes, Sarah; Batistatou, Evridiki; Camacho, Elizabeth; Knowles, Sarah; Darley, Sarah; Grande, Gunn; Ewing, Gail; Bowen, Audrey

VERSION 1 – REVIEW

REVIEWER	Monique R. Pappadis University of Texas Medical Branch at Galveston, USA
REVIEW RETURNED	19-Apr-2020

GENERAL COMMENTS	Overall, this study aimed to reduce caregiver strain using the CSNAT-Stroke intervention compared to usual care. In addition, the clinical and cost-effectiveness was assessed. No significant differences between the intervention and control groups were identified. However, the authors do not provide details into possible reasons for the lack of differences, nor do they provide possible modifications to improve clinical effectiveness or improve cost effectiveness compared to usual care. Authors suggest this study to be a pragmatic trial; however, one could argue that it is an explanatory trial instead of a pragmatic trial, given the need for extensive training of the intervention. Abstract Please indicate by whom were the carers referred. Indicate that UK voluntary 'sector stroke specialist organisation' made the referrals. The data for the economic evaluation include more than self-reported healthcare utilisation and QoL, correct? If so, please include. Was the outcome assessor blinded? Lines 40-45: Language could be tighter and some details can be moved to the methods/results sections. Repetitious: the intervention was not implemented fully – it is repeated twice in the abstract. Consider revising the strengths and limitations of the study. It should not be a restatement of the findings, but a summary of how this work adds to the literature and any important limitations to be considered in light of the results. Introduction Page 3, Lines 46-47. The beneficial use of the CSNAT could be better demonstrated by providing a brief summary of its positive

	outcomes. In addition, more detail on the intervention would be appreciated. Lines 54-57: Greater details on the make-up of the carer advisory research group would be helpful. Methods/Results Page 5, Lines 46-47: were the outcome assessors blinded? Although the protocol was included, greater details are suggested regarding the randomization (why were 2 blocks chosen?), the intervention, the procedures, details regarding the coordinators/service providers, information regarding treatment fidelity, reporting cluster-level data in addition to individual-level data, etc. Not much emphasis was placed on the information regarding the service providers, and information regarding the stroke severity of the carers' loved ones with stroke. In addition, the details regarding the intervention are rather vague, and this study as presented would be challenging to replicate, even after reviewing the published protocol. Discussion Have the authors also considered that the outcome measure chosen may not be sensitive enough to detect changes and the social desirability component? Given that implementation challenges have been previously noted, it would be nice if the authors include a greater discussion on what those challenges were, were the same challenges identified in this study, how did the researchers attempt to address the challenges prior to study implementation, as well as provide suggestions for future work using the CSNAT.

REVIEWER	Shinta Nishioka Nagasaki Rehabilitation Hospital, Japan
REVIEW RETURNED	25-Apr-2020

GENERAL COMMENTS	This study investigated an effect of new strategy (NSNAT-Stroke) to support the carer of stroke survivors. The authors found that CSNAT-Stroke showed no advantage over the standard care with slightly increase in cost, which may be due to insufficient implementation of the intervention. Study background and aims care clear and concise. Study protocol are very rigorous. Results are accurately described according to the protocol and the statistical analyses were rigorously conducted. Discussions are well-organized and there are no extended interpretation of the results. I think this study successfully added the scientific value for supporting caregiver of stroke patients, and will pave the way to future research. I suggest that the author address a few minor points. #1 Were the study participants caring for the stroke survivors alone? Were they caring in collaboration with other carer (e.g., other family members)? I think sharing care burden with other person can attenuate the caregivers' strain. In addition, I'm interested in what kind of social support stroke patients were

	receiving, which can also affect the burden of caregiver. Please add these information in the Results, if applicable. #2 P12: “These findings were consistent across all sensitivity analyses including: excluding delayed responders; removing carer dyads; imputing missing outcome data; and combining 3 and 6 month data, suggesting that the results are robust to assumptions made in the analysis.” I would like to know the results of per protocol analysis. Adding the Table of sensitivity analysis for primary and secondary outcome may be useful for readers.

REVIEWER	Dr Jack Parker University of Derby, UK
REVIEW RETURNED	30-Apr-2020

GENERAL COMMENTS	This is an interesting and much needed area of research. The paper is presented and written well with excellent detail in places. Essentially, this paper could include a list of recommendations for – if we were to do this again, we would… 1. I would like to have read more regarding the specific contribution of the PCPI – what were there recommendations? How did you alter the trial based on these? 2. The paper would benefit form more narrative reflection on the recruitment and limitations of recruitment (i.e. if you were to recruit again, what would you do differently), the index of multiple deprivation (IMD) rank of the participants would also be interesting to explore. 3. What recommendations would the authors suggest in terms of outcome measures, data collection points, procedures, clinician involvement. 4. The paper would benefit from some reflection of the study design and its strength and limitations for a complex intervention for a complex condition with reference to the MRC framework.

REVIEWER	Obioha Ukoumunne University of Exeter, UK
REVIEW RETURNED	18-Sep-2020

GENERAL COMMENTS	The paper summarises the findings a of cluster randomised trial of the effectiveness and cost-effectiveness of an intervention for providing support to carers of people who have had a stroke. The paper is very clearly written (adhering pretty strictly to the CONSORT statement) and the study is well designed and analysed. I have some minor comments: 1. I think it would be good to add the mean and standard deviation of the primary outcome for each trial arm to the Abstract, just to help contextualise the raw effect. 2. In the Study Design section (3rd line) I would add a sentence stating that stroke specialist provider services (clusters) were randomised. This is stated in the Abstract but I think it’s worth saying in the main text also. 3. At the start of the second paragraph in the “Study Design” section, maybe change “This paper focuses on the cRCT’s …..” to “This paper focuses on the intervention’s….”.

	4. The staff of the specialist stroke providers were not masked to trial arm status. Even though they were responsible for recruiting participants within clusters there does not appear to be any bias that has resulted based on the number of participants in each arm, which is good. I just wondered if the authors did anything specific to make sure all eligible people were approached consistently. 5. In the first bullet point in the "Outcomes" section it says in relation to the FACQ measure that "Each subscale produced a mean score out of five…". Does this mean from a possible 1 to 5 or 0 to 5? 6. The authors fitted generalised linear models specifying the gamma distribution and log link function for analysing cost data. What approach was used to allow for clustering in this analysis? 7. At the beginning of the Results section it says that 39 clusters were randomised and 4 clusters withdrew leaving 32 clusters. Should this not read 35 clusters (i.e., 39 minus 4) rather than 32? The same paragraph then says an extra three clusters were randomised which would take the number up to 38? I think a bit more clarity is needed in these sentences. 8. In the Results section when reporting the interquartile range for time from stroke event to initiation of support, I would show the lower and upper bounds for this range rather than the difference between them. 9. The full details of the sample size calculation are provided in the protocol paper but I do think this paper would benefit form a short description of the calculation, providing enough detail so that it can be replicated.

VERSION 1 – AUTHOR RESPONSE

R1

Overall, this study aimed to reduce caregiver strain using the CSNAT-Stroke intervention compared to usual care. In addition, the clinical and cost-effectiveness was assessed. No significant differences between the intervention and control groups were identified. However, the authors do not provide details into possible reasons for the lack of differences, nor do they provide possible modifications to improve clinical effectiveness or improve cost effectiveness compared to usual care.

- We are surprised by this statement about lack of detail into reasons and modifications as we have several sections on these issues in the discussion (e.g. pages 15 & 16 under 'strengths..implications'). We hope we have addressed reviewer 1's concerns by adding the following signposting to the opening paragraph of the discussion (page 15):
"There are several possible explanations for our neutral finding explored in detail below."

In addition to the methodological modifications discussed (e.g. choice of comparator, outcome measures) we report that full implementation of the intended intervention was not achieved, the implication being that if it were then perhaps it would improve effectiveness. On page 16, para 5, we also point to our process evaluation paper (revisions submitted to BMJ Open) for a detailed exploration of our evidence for these conclusions.

Authors suggest this study to be a pragmatic trial; however, one could argue that it is an explanatory trial instead of a pragmatic trial, given the need for extensive training of the intervention.

- We are aware that few trials are either fully pragmatic or fully explanatory. OSCARSS was designed as a pragmatic trial in that it was powered to evaluate the effectiveness of a new intervention in real-life conditions (service users, providers and settings). Staff training was not extensive and was delivered in collaboration with the provider organisation's training department. An explanatory trial would have been designed differently e.g. with researchers delivering the intervention in idealised conditions/settings.

Abstract: Please indicate by whom were the carers referred. Indicate that UK voluntary 'sector stroke specialist organisation' made the referrals

- We have reworded the abstract to say:

"Setting: Clusters were services, from a UK voluntary sector stroke specialist provider, delivering support primarily in the homes of both stroke survivors and their informal carers.
Participants: Adult carers in participating clusters were referred to the study by cluster staff following initial support contact."

Abstract: The data for the economic evaluation include more than self-reported healthcare utilisation and QoL, correct? If so, please include

- Thank you, we have expanded this to say:

"The economic evaluation included self-reported healthcare utilisation, intervention costs, and EQ-5D-5L."

Abstract: Was the outcome assessor blinded?

- We prefer the term masked and have revised this to say:

"Carer research participants who provided self-reported outcome data were unaware of allocation"

Abstract: Lines 40-45: Language could be tighter and some details can be moved to the methods/results sections.

- Lines 40-45 relates to the secondary analyses etc in the abstract's results section. We are reluctant to remove any details from the abstract but welcome the editor's direction.

Abstract Repetitious: the intervention was not implemented fully – it is repeated twice in the abstract

- Thank you, we edited the phrase in 'results' to read "intervention fidelity was not achieved" to reduce the repetition from the conclusion.

Consider revising the strengths and limitations of the study. It should not be a restatement of the findings, but a summary of how this work adds to the literature and any important limitations to be considered in light of the results.

- Revised to say:

"• We successfully conducted the first adequately-powered cluster randomised controlled trial of an approach to support informal carers of stroke survivors, but may have benefitted from a feasibility trial to maximise intervention fidelity.
• We collaborated closely with service providers and previous service users to pragmatically tailor the intervention for implementation, including a staff training package.
• The demographic profile of the sample was as expected for carers of stroke survivors but the sample lacked ethnic diversity and we may have benefitted from seeking data beyond six months after support had been initiated.
• We highlight the feasibility of robust research with this population and signpost to suggestions from our nested process evaluation for improved implementation of person-centred care."

Page 3

Introduction
Page 3, Lines 46-47. The beneficial use of the CSNAT could be better demonstrated by providing a brief summary of its positive outcomes.

In addition, more detail on the intervention would be appreciated.

- Thank you for the suggestions. Revised to say:

"It was developed, implemented and tested in the context of palliative care with positive outcomes, including a significant reduction in caregiver strain as measured on the Family Appraisal of Caregiving Questionnaire (FACQ)."

And on page 4:
"In close collaboration with a study-specific carer advisory research group (see PPI below) and a UK stroke service provider organisation, we adapted the CSNAT intervention including a staff training and implementation package tailored to the provider organisation (see Interventions below, Figure 1, Table 1 and Table S1 supplementary material)."

Introduction
Lines 54-57: Greater details on the make-up of the carer advisory research group would be helpful.

- We have signposted to the PPI section and revised the text to say:

"A video summarising their role in OSCARSS is available on the study website: https://www.arc-gm.nihr.ac.uk/projects/oscarss and we have published a separate paper on the working practices and

experiences of RUG members and the researchers who facilitated the group meetings, following the GRIPP2 framework."

Methods/Results
Page 5, Lines 46-47: were the outcome assessors blinded?

- We prefer the term masking to blinded and have revised our 'randomisation and masking' section to say:

"Carer research participants provided self-report primary and secondary outcomes unaware of allocation; they received support from their local randomised cluster and consented to follow-up data collection only. Carers were told that the service was being evaluated but not told about the randomised clusters." (page 6)

Although the protocol was included, greater details are suggested regarding the randomization (why were 2 blocks chosen?), the intervention, the procedures, details regarding the coordinators/service providers, information regarding treatment fidelity, reporting cluster-level data in addition to individual-level data, etc. Not much emphasis was placed on the information regarding the service providers, and information regarding the stroke severity of the carers' loved ones with stroke.

- Thank you for the suggestion.

Blocks of size 2 were chosen to ensure that the two groups were similar in terms of the number of clusters. Note that all initial clusters were randomised simultaneously so we were not at any risk of bias from researchers being able to predict the next allocation (as would be the case in an in individually randomised trial). We have edited the 'randomisation and masking' section to give a justification of the block sizes, which now reads:

"Clusters were block randomised to intervention or control, with stratification for size of service using random blocks of two (to ensure similar numbers of carers and clusters in each arm). The trial statistician performed the randomisation of all recruited clusters simultaneously using an anonymised list of cluster ID numbers and size of service data." (Page 5)

We have included some of the additional details in our supplementary information and added a table on cluster characteristics to supplementary information and have added a signposting line in the first paragraph of the results section: "Cluster and staff baseline characteristics are included in supplementary materials, Table S2)" (page 9)

To keep the word count of this paper manageable, we signpost to the open access protocol paper. Intervention fidelity, and information regarding the service providers, is reported in our signposted process evaluation paper. PLEASE NOTE: these papers are 'sister publications' both currently under review by BMJ Open. We welcome the opportunity to include full references from each paper to the other.
However if the editor would like additional tables for the current paper please let us know.

As carers are the participants in this study stroke severity is reported as the carer's perception of independence 'amount of care needed', as opposed to therapist-observed impairment severity (page 10).

In addition, the details regarding the intervention are rather vague, and this study as presented would be challenging to replicate, even after reviewing the published protocol.

- We recognise this is a common challenge when reporting trials of complex interventions. As the current paper focuses on the trial and economic evaluation we have aimed to provide all necessary detail for methodological replication. We refer the reviewer to the detailed TIDIER table in our supplementary data and to our process evaluation sister paper to replicate the intervention itself. We hope the editor agrees with this approach.

Discussion
Have the authors also considered that the outcome measure chosen may not be sensitive enough to detect changes and the social desirability component?

- We have added this to page 17 as follows:

"Whilst our choice of primary outcome was informed by past research using the CSNAT intervention and the preferences of our service user group of stroke carers, our measure may not have been adequate to detect a difference in our population of stroke carers."

Discussion
Given that implementation challenges have been previously noted, it would be nice if the authors include a greater discussion on what those challenges were, were the same challenges identified in this study, how did the researchers attempt to address the challenges prior to study implementation, as well as provide suggestions for future work using the CSNAT.

- We thank the reviewer for this observation and agree that the nuance of implementation is a topic of interest. This is discussed more in our sister Process Evaluation and we have added another signpost to this paper in our discussion (page 15) which says:

"Several UK studies by the CSNAT team showed similar outcomes and good acceptability, but also reported implementation challenges similar to those found in OSCARSS, and discussed in our sister process evaluation paper".

We ask the BMJ Open editors to coordinate publication of these two papers and, in doing so, allow us to update the reference lists of each to be able to appropriately link with references.

We have strictly followed CONSORT and BMJ Open guidance for this paper content and are in line with recommended word limits so feel it is important to signpost where appropriate rather than duplicate.

Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014;348:g1687.

R2

Study background and aims care clear and concise. Study protocol are very rigorous. Results are accurately described according to the protocol and the statistical analyses were rigorously conducted. Discussions are well-organized and there are no extended interpretation of the results.
I think this study successfully added the scientific value for supporting caregiver of stroke patients, and will pave the way to future research.
I suggest that the author address a few minor points.

- We thank reviewer 2 for this positive feedback

#1 Were the study participants caring for the stroke survivors alone? Were they caring in collaboration with other carer (e.g., other family members)? I think sharing care burden with other person can attenuate the caregivers' strain. In addition, I'm interested in what kind of social support stroke patients were receiving, which can also affect the burden of caregiver. Please add these information in the Results, if applicable.

- We thank the reviewer for this observation but, we did not collect these specific data so cannot add to results. We have added a sentence to Methods 'participants' section (page 5) and one to Discussion (page 10) to acknowledge this as follows:

"We aimed to recruit those individuals identified as 'primary caregiver', even when there may have been other informal carers involved." Page 5

"We aimed to recruit the primary caregiver but did not collect additional data on whether they were caring alone or with support." Page 10

#2 P12: "These findings were consistent across all sensitivity analyses including: excluding delayed responders; removing carer dyads; imputing missing outcome data; and combining 3 and 6 month data, suggesting that the results are robust to assumptions made in the analysis." I would like to know the results of per protocol analysis. Adding the Table of sensitivity analysis for primary and secondary outcome may be useful for readers.

- Thank you for the suggestion but given that the results of the sensitivity analyses were consistent with the main analyses, and we are at capacity with the number of tables, we don't believe that providing that level of extra detail would be useful. We welcome the editor's decision on that.

The per protocol analysis that we did was planned in the Statistical Analyais Plan as excluding clusters or participants who deviated from the protocol according to certain criteria. The only protocol deviations we had, according to our original definitions, were participants who returned outcomes data outside of intended date windows, and therefore this this analysis changed our estimates very little. The reviewers comment has made us realise that the term 'per protocol analysis' is perhaps unclear as it might be expected that this would exclude participants who did not receive the intervention as planned - unfortunately we had no measure of this and from the outset we were interested in the pragmatic effect of implementing the intervention rather than the complier average causal effect.

We have amended our description of the sensitivity analysis in the 'Statistical and Economic Analysis' section to state: "Sensitivity analyses were pre-specified in the Statistical Analysis Plan to explore any potential bias and examine the robustness of findings"
Page 8

R3
This is an interesting and much needed area of research. The paper is presented and written well with excellent detail in places. Essentially, this paper could include a list of recommendations for – if we were to do this again, we would…

- We thank Reviewer 3 for their endorsement of this research and their positive feedback on presentation style.

We have followed reporting guidelines and included strengths and limitations of the methodology in our discussion and we addresses reviewer 3's specific recommendations in subsequent comments.

---
1. I would like to have read more regarding the specific contribution of the PCPI – what were there recommendations? How did you alter the trial based on these?

- We thank the reviewer for their comments which value the PPI contribution. Word limits don't permit us to add much more detail but we have signposted to a separate published paper where we discuss in detail the contribution and recommendations made by PPI members and edited the text on page 5 so it now says:

"The RUG advised on participant recruitment and were central in limiting the burden of participation for carers. The RUG also were key in supporting adaptation of the research intervention (CSNAT-Stroke) and staff training package, including role-playing videos of the intervention in practice. A video summarising their role in OSCARSS is available on the study website: https://www.arc-gm.nihr.ac.uk/projects/oscarss and, following GRIPP2 framework18 we have published a separate paper on the working practices and experiences of RUG members and the researchers who facilitated the group meetings16"

In line with feedback recently received from our sister publication (process evaluation), we have also changed the header of this section from "Patient and Carer Public Involvement" to "Patient and Public Involvement" (page 4)

---
2. The paper would benefit from more narrative reflection on the recruitment and limitations of recruitment (i.e. if you were to recruit again, what would you do differently), the index of multiple deprivation (IMD) rank of the participants would also be interesting to explore.

- We thank the reviewer for this comment. We would like to highlight that 66% consented to participate and that we feel this is a realistic figure for this early timepoint after stroke when many people have not yet self-identified as carers.
As such, we would not recruit differently; if we recruited at a later timepoint we may have increased our response rate but such a delay would have answered a different research question and compromised our belief that initiating the intervention early may be beneficial for carers.
We did not collect IMD rank of participants and believe that post-hoc exploration of this would not add to the paper but we thank the reviewer for their interest.
No changes have been made to the paper in light of the above.

---
3. What recommendations would the authors suggest in terms of outcome measures, data collection points, procedures, clinician involvement.

- We have confidence in our findings for this population and are reticent to make recommendations for methodological changes in this paper which focuses on the trial and economic evaluation. OSCARSS' main recommendations are more around addressing the challenges of fully implementing person-centred care in research and service development, including greater attention to change management process. These are covered in detail in the sister process evaluation paper (e.g. champions, managerial ownership) and briefly mentioned in the conclusions of the current paper (pages 16-17) as follows:

".. need to be addressed through enhanced and ongoing staff training as well as organisational mechanisms to support and champion new approaches becoming embedded into practice. There remains a high priority for research to determine how best to support carers of stroke survivors."

---
3. What recommendations would the authors suggest in terms of outcome measures, data collection points, procedures, clinician involvement.

- We have confidence in our findings for this population and are reticent to make recommendations for methodological changes in this paper which focuses on the trial and economic evaluation. OSCARSS' main recommendations are more around addressing the challenges of fully implementing person-centred care in research and service development, including greater attention to change management process. These are covered in detail in the sister process evaluation paper (e.g. champions, managerial ownership) and briefly mentioned in the conclusions of the current paper (pages 16-17) as follows:

".. need to be addressed through enhanced and ongoing staff training as well as organisational mechanisms to support and champion new approaches becoming embedded into practice. There remains a high priority for research to determine how best to support carers of stroke survivors."

---
4. The paper would benefit from some reflection of the study design and its strength and limitations for a complex intervention for a complex condition with reference to the MRC framework.

- We have revised the strengths and limitations (page 3) to say:

"We successfully conducted the first adequately-powered cluster randomised controlled trial of an approach to support informal carers of stroke survivors, but may have benefitted from a feasibility trial to maximise intervention fidelity."

R4
The paper is very clearly written (adhering pretty strictly to the CONSORT statement) and the study is well designed and analysed.

- We thank reviewer 4 for these positive comments and acknowledgement of our adherence to guidance when reporting our findings.

---

1. I think it would be good to add the mean and standard deviation of the primary outcome for each trial arm to the Abstract, just to help contextualise the raw effect.

- We felt it was more useful to provide the adjusted mean difference and confidence interval in the abstract. We do report the mean and SDs of the primary (and secondary) outcomes per arm in Table 3 pages 11-12.

Current edits have taken the abstract to 400 words and more words would be required to add these details to the abstract. We would value the editor's feedback on this decision.

To aid the decision, adding these details would read as follows and make the abstract word count = 424:
Primary outcome measure: intention to treat analysis for 84% retained participants (175 intervention; 174 control) found mean (SD) FACQ carer strain at 3 months to be 3·11 (0·87) in the control group compared to 3·03 (0·90) in the intervention group, adjusted mean difference of -0·04 (95% CI -0·20 to 0·13).

---

2. In the Study Design section (3rd line) I would add a sentence stating that stroke specialist provider services (clusters) were randomised. This is stated in the Abstract but I think it's worth saying in the main text also.

- We do say who the clusters were, starting from the 3rd line, as follows:

"Clusters were drawn from services commissioned by NHS or Local Authorities, delivered by a UK voluntary sector stroke specialist organisation providing long-term support to stroke survivors and carers, including hospital and home visits."

We use the word randomised twice in the previous 2 sentences.

We would value the editor's feedback on any wording required.

---

3. At the start of the second paragraph in the "Study Design" section, maybe change "This paper focuses on the cRCT's ….." to "This paper focuses on the intervention's….".

- Thank you, we have revised this sentence to:

"This paper focuses on the RCT to explore the intervention's clinical and cost-effectiveness."

---

4. The staff of the specialist stroke providers were not masked to trial arm status. Even though they were responsible for recruiting participants within clusters there does not appear to be any bias that has resulted based on the number of participants in each arm, which is good. I just wondered if the authors did anything specific to make sure all eligible people were approached consistently.

- This is a good point. We have added a line to the end of the 'participants' section (page 5) as follows:

"Researchers were in regular contact with all cluster staff and senior leadership to encourage fidelity with research procedures, including the consistent invitation to participate for all eligible carers."

---

5. In the first bullet point in the "Outcomes" section it says in relation to the FACQ measure that "Each subscale produced a mean score out of five…". Does this mean from a possible 1 to 5 or 0 to 5?

- Thank you. We have added clarifying text on page 7 so it now reads as:

"Each item was scored from 1 to 5 and each subscale produced a mean score out of five, with a score of three as neutral, and higher scores indicating a greater amount of the variable being measured."

---

6. The authors fitted generalised linear models specifying the gamma distribution and log link function for analysing cost data. What approach was used to allow for clustering in this analysis?

- We have revised the text at the top of page 9 to say:

"Models were adjusted for clustering using the same covariates as the clinical-effectiveness analysis."

---

7. At the beginning of the Results section it says that 39 clusters were randomised and 4 clusters withdrew leaving 32 clusters. Should this not read 35 clusters (i.e., 39 minus 4) rather than 32? The same paragraph then says an extra three clusters were randomised which would take the number up to 38? I think a bit more clarity is needed in these sentences.

- We are grateful to the reviewer for spotting our mistake and have corrected the CONSORT flowchart (fig 2) and text on page 9 as follows:

"In September 2016 we randomised 36 clusters (18 intervention; 18 18 control). Three control and one intervention cluster withdrew soon after due to decommissioning or all staff long-term sickness (see Figure 2) so 32 clusters were trained in January 2017. Three replacement clusters were recruited, randomised and trained between February and April 2017 (one intervention; two control). This gave a total of 35 recruiting clusters (18 intervention; 17 control)."

PLEASE NOTE: our Figure 2 as uploaded to the BMJ Open system keeps being changed to inverted colours (black background when it should be white). We have tried uploading different versions of the figure but it keeps getting changed. We would value support to ensure the correct Figure is included in the published manuscript

---

8. In the Results section when reporting the interquartile range for time from stroke event to initiation of support, I would show the lower and upper bounds for this range rather than the difference between them.

- Thank you for this comment. We have amended text in the results section to read:

"The mean age of carers was 62 years old when they joined the study and the median time from the stroke event to support being initiated was 2·3 months across the whole sample (IQR = 1.1 to 2.3)."

In line with this recommendation, we have also adjusted the IQR displayed in Table 2 for Independence of stroke survivor, as rated by the carer. IQRs now read "(8 to 14)" rather than "(6)" for each arm
Page 10-11

---

9. The full details of the sample size calculation are provided in the protocol paper but I do think this paper would benefit form a short description of the calculation, providing enough detail so that it can be replicated.

- We thank this reviewer. We missed out the word 'standardised' and the significance level from our description of the sample size – we now believe that all necessary parameters to recreate the sample size calculation are included. We have also added reference to the STATA package used so the 'Statistical and Economic Analysis' now reads:

"A minimum of 400 carers recruited from 32 clusters (200 per trial arm) would provide 80% power to detect standardised effect sizes on the primary outcome of 0·31 or more (FACQ Strain mean score), assuming an intraclass correlation coefficient of 0.01 with a 20% loss to follow-up, at the 5% significance level. Power was calculated using the STATA clsampsi function23."
Page 8

New Ref 23: Batistatou E, Roberts C, Roberts S. Sample size and power calculations for trials and quasi-experimental studies with clustering. The Stata Journal. 2014;14(1):159-75.

In addition to the Editor and Reviewer comments there are minor additions to address the Editorial Office feedback received on 4th November 2020 after submission of the files for consideration, namely:

1. Edit to contributor statement to correct an initial from "SR" to "SD" (page 17)
2. Funder statements revised to indicate 'N/A' for grant number (page 18)
3. Citing all supplementary file tables "(Tables S3 to S8)" within the main manuscript itself. (page 13)

VERSION 2 – REVIEW

REVIEWER	Monique R. Pappadis, MEd, PhD University of Texas Medical Branch at Galveston, USA
REVIEW RETURNED	29-Nov-2020

GENERAL COMMENTS	Thank you for addressing the reviewers' concerns. This manuscript is well-written and provides a contribution to the literature. Minor: Figure 2 is challenging to read with the black background. It is possible that the copyedited version would be fine.

REVIEWER	Shinta Nishioka Nagasaki Rehabilitation Hospital, Japan
REVIEW RETURNED	24-Nov-2020

GENERAL COMMENTS	I have no further comment. This manuscript is now suitable for publication.

REVIEWER	Obi Ukoumunne NIHR ARC South West Peninsula, University of Exeter, UK
REVIEW RETURNED	28-Nov-2020

GENERAL COMMENTS	I have some minor remaining comments. 1. The authors say that they allowed for clustering, but they have still not said what method was used to allow for clustering. They have fitted generalised linear models to analyse the data, but the ordinary version of these models do not allow for clustering. What method was used to extend the GLM so that it allows for clustering in the data? 2. The authors should add to the abstract the mean and standard deviation for the FACQ for each of the intervention and control groups. The standard deviation, particularly, is important so that the reader can make a judgement about whether or not the 95% confidence interval for the adjusted mean difference includes differences that would be clinically important if true. It is relevant here because the authors describe their 95% confidence interval as "tight".

VERSION 2 – AUTHOR RESPONSE

R2	I have no further comment. This manuscript is now suitable for publication.	Thank you. We are very happy to hear this.
R4	I have some minor remaining comments.	We agree the below outstanding comments are minor and we thank reviewer 4 for raising them. We address them below.
R4	1. The authors say that they allowed for clustering, but they have still not said what method was used to allow for clustering. They have fitted generalised linear models to analyse the data, but the ordinary version of these models do not allow for clustering. What method was used to extend the GLM so that it allows for clustering in the data?	Clustering was accounted for in 2 ways. Cluster-level covariates (size of service; and experience of staff delivering support) were included the model (as was done in the clinical effectiveness models) and the model was specified so that the 95% CI allowed for intragroup correlation using the vce(cluster) option in Stata. We have adjusted the text on page 8 to read: "Models allowed for clustering by adjusting for the same cluster-level covariates as the clinical-effectiveness analysis (see above) and the models were specified so that the confidence intervals allowed for intragroup correlation."

R4	2. The authors should add to the abstract the mean and standard deviation for the FACQ for each of the intervention and control groups. The standard deviation, particularly, is important so that the reader can make a judgement about whether or not the 95% confidence interval for the adjusted mean difference includes differences that would be clinically important if true. It is relevant here because the authors describe their 95% confidence interval as "tight".	We have amended the abstract text to read: "Primary outcome measure: intention to treat analysis for 84% retained participants (175 intervention; 174 control) found mean (SD) FACQ carer strain at 3 months to be 3·11 (0·87) in the control group compared to 3·03 (0·90) in the intervention group, adjusted mean difference of -0·04 (95% CI -0·20 to 0·13)" Please note that this takes our abstract to 424 words so some text has been cut off in the online submission system
R1	Thank you for addressing the reviewers' concerns. This manuscript is well-written and provides a contribution to the literature.	Thank you. We are very happy to hear this.
R1	Minor: Figure 2 is challenging to read with the black background. It is possible that the copyedited version would be fine.	We have uploaded various versions of Figure 2 with the white background and whilst they all look correct initially, they subsequently seem to 'corrupt' to have a black background. We have attempted to re-loaded the Figure again but if it corrupts later, we are sorry we cannot fix this but look forward to working with the copyeditor to do so.

We note that we have no comments from Reviewer 3.

In addition to the above comments there are minor additions to address the Editorial Office feedback received on 5th December 2020 after submission of the files for consideration, namely:

1. Edit to author initial from "SR" to "SAR" to reflect their entry in ScholarOne system. This affects the author list (page 1) and contributor statement (page 17 and ScholarOne system)
2. Updating our 'CONSORT Cluster Trials Checklist' file upload with page numbers. PLEASE NOTE: We have used page numbers from the marked copy of the 'Main Document' and PDF supplementary document.

VERSION 3 – REVIEW

REVIEWER	Obi Ukoumunne University of Exeter, UK
REVIEW RETURNED	10-Dec-2020

GENERAL COMMENTS	All my previous comments have been well addressed.