PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	A randomised controlled trial of the clinical and cost-effectiveness of a peer delivered self-management intervention to prevent relapse in crisis resolution team users: study protocol
AUTHORS	Johnson, Sonia; Mason, Oliver; Osborn, David; Milton, Alyssa; Henderson, Claire; Marston, Louise; Ambler, Gareth; Hunter, Rachael; Pilling, Stephen; Morant, Nicola; Gray, Richard; Weaver, Tim; Nolan, Fiona; Lloyd-Evans, Brynmor

VERSION 1 – REVIEW

REVIEWER	Matthijs Blankers Arkin Mental Health Care - The Netherlands
REVIEW RETURNED	30-May-2017

GENERAL COMMENTS	Dear authors, I have read your protocol ms with great interest. It is clearly written and well structured. Nice work. I have a number of comments and suggestions you might want to consider while revising your paper. p3. Peer supporters have themselves experiences ill mental health: but they do not nescessarily have experienced psychiatric crisis? Why did you make this choice? Wouldnt it be better to select peer workers with a history of psychiatric crisis? p4. Hypotheses - Please quantify the hypotheses. For your power analysis you indicate that you expect 50% vs 35% admission rates at 1yr follow up, this should also be mentioned in your hypotheses 1. - For hypothesis 2: how much longer / fewer days? - hypothesis 4 is not an actual hypothesis. Either state what you expect to find or report this planned exploratory analysis elsewhere in the ms. p5. - what/who determines which patients will be seen by a CRT team? How does this group compare to all crisis care patients? In/exclusion criteria - Why did you omit criteria related to age (both underage and adults will be included?), type of diagnosis, first/continued CRT contact? Doesn't this make your sample very heterogeneous? - How will you assess capacity to give IC in this population?

	- Patients who are to unstable/unsave to visit at home, do you exclude them / how do you manage the safety of the peer workers? - You indicate (although not a formal in/excl. criterion) that >=50% of patients should have psychotic or bipolar disorder as a diagnosis. How will you manage this? Will you introduce this diagnosis requirement as an in/excl. criterion if you do not meet the >=50% criterion? Please elaborate. p9. Instruments - Why do you use the (now outdated/old) 3 level version of the EQ5D? - Recovery promoting relationships scale: is this feasible to measure therapeutic alliance when patients see multiple peer workers? Will the instrument focus on the therapeutic alliance with one peer worker of is it a more general measure of alliance? - Have you considered to measure the working alliance by interviewing the peer workers as well? p10. - Why did you choose a 1:2 allocation rate? - Where is the (relatively low) ICC of .03 based on? p11. - changes to the procedures were sufficiently minimal: what were the changes? - Have you considered to use stratified randomisation by centre or diagnosis? Why (not)? - Will you register characteristics of patients who choose not to participate in the study? p14. - You indicate you will use complete case analysis and refer to this as ITT. However complete case analysis is not ITT. For ITT missing data should be addressed adequately, eg through multiple imputation or through (G)LMM with maximum likelihood estimation, and with variables associated to non-reponse in the model. Please consider adjusting this section accordingly. - is the cluster size of 1 appropriate? shouldn;t it be 1 cluster for all participants? - "random effects": do you plan a random slope, random intercept, or both? Please elabore - Results will be reported as OR while before you mentioned percentages (50% vs 35%). Please consider harmonising this. p15. Linear modeling is not the most appropriate for count data (days spent in care), could you address this? p16. Please define the acronym PSW. Is there 1 PSW per patient or multiple. I had the impression multiple but I may be wrong. If 1 per patient: how are patient and PSW matched/paired? p17. Economic evaluation. It is somewhat more common to have the societal perspective analysis as the base care. Please refer to the ISPOR 2015 guidelines or see Drummond et al 2015. p18 - first line: 'costed' -> 'valued' - Confidence intervals: are these 95% intervals? - it is somewhat more common to report median ICERS (instead of

	mean) - how will you handle ICERs in case of bootstrapped samples in the dominated quadrant (less effects, more costs)? - Are other health care contacts (except mental health and GP) also recorded and used in the analysis? For discussion / general - It seems that the control group only receives an information booklet while the PSW group receives care which may be considered 'normal care': making a crisis/emergency plan etc. Will the control group receive the actual care as ususal in the service area where the study is performed or is the usual care situation more like the care the PSW group receives? - CRT workers are blinded to the study condition the patient is in, but what about other (mental) health workers? May this have an impact on your study? - Please include a discussion/strengths/limitations section near the end of the ms.

VERSION 1 – AUTHOR RESPONSE

Reviewer's comments

I have read your protocol ms with great interest. It is clearly written and well structured. Nice work. I have a number of comments and suggestions you might want to consider while revising your paper.

We are grateful for the kind comment on the quality of the paper!

p3. Peer supporters have themselves experiences ill mental health: but they do not nescessarily have experienced psychiatric crisis? Why did you make this choice? Wouldnt it be better to select peer workers with a history of psychiatric crisis?
Our criterion for recruitment of Peer Support Workers was personal experience of mental health problems: this is an agreed essential requirement for a mental health peer support role (Davidson et al. *2006; Mead and MacNeil, 2006) , and most people with personal experience of mental health problems are likely to have experienced a crisis of some type. We did not require personal experience of being on the caseload of a crisis team – this would have potentially created delays in recruitment and in conducting the trial on time, and was consistent with our decision not to attempt to match Peer Support Workers and participants beyond the essential shared experience of mental health difficulties. We have further clarified this in the section on "Peer Support Workers and their Training" (pp 8-9).

p4. Hypotheses
- Please quantify the hypotheses. For your power analysis you indicate that you expect 50% vs 35%
We have now added in the section on the primary hypothesis the size of the difference which was the basis for our power calculation (p5).

- For hypothesis 2: how much longer / fewer days?

The secondary hypotheses (Hypothesis 2) were not the basis for our study power calculation, so the anticipated differences have not been quantified.

- hypothesis 4 is not an actual hypothesis. Either state what you expect to find or report this planned exploratory analysis elsewhere in the ms.
We agree – we have amended this so that this exploratory analysis no longer appears on the numbered list of hypotheses (p5).

p5.
- what/who determines which patients will be seen by a CRT team? How does this group compare to all crisis care patients?
We have now added to the Background and Rationale section a brief description of the CRT target group and a further reference describing this (p3).
In/exclusion criteria
- Why did you omit criteria related to age (both underage and adults will be included?), type of diagnosis, first/continued CRT contact? Doesn't this make your sample very heterogeneous?

We accepted a mixed group of CRT service users because we wished to develop and test an intervention sufficiently broad and flexible to be potentially effective in this population as a whole. Thus service use rather than clinical characteristics defines the sample. We have added a note to clarify this in the section on inclusion and exclusion criteria (p6).

- How will you assess capacity to give IC in this population?

The recruitment process begins with CRT clinicians approaching potential participants to ask if they would be happy to be contacted by a researcher to discuss the trial. Initial assessment of capacity to give informed consent to enter the trial is by these clinicians: they approach only services users whom they assessed as having such capacity. Research staff contacting potential participants then make a further assessment of capacity. They are provided with training about assessing capacity, following Royal College of Nursing guidance, and base their assessment on discussion of the study information sheet with potential participants. Researchers are directed not to recruit any participants about whose capacity they had concerns, but to discuss them further with clinician and exclude them if capacity is unclear. Capacity is reassessed at each follow-up interview. A summary of this process has now been inserted in the Participant Timeline section (p11).

- Patients who are to unstable/unsafe to visit at home, do you exclude them / how do you manage the safety of the peer workers?

Local Trust lone worker policies are adhered to and risk assessments discussed with local clinicians. Precautions include seeing service users on NHS premises if they were felt to be too risky to see at home, and checking on the safety of both peer support workers and researchers immediately following each visit. A note explaining this is now included in the section on the peer support worker intervention (p9).

- You indicate (although not a formal in/excl. criterion) that >=50% of patients should have psychotic or bipolar disorder as a diagnosis. How will you manage this? Will you introduce this diagnosis requirement as an in/excl. criterion if you do not meet the >=50% criterion? Please elaborate.

The number of people with such a diagnosis is monitored at each site throughout the recruitment period, If sites have recruited 50% of their intended number who do not have a psychotic bipolar diagnosis, they subsequently recruit only people with such a diagnosis.

p9. Instruments
- Why do you use the (now outdated/old) 3 level version of the EQ5D?

This was current at the time the study was designed.

- Recovery promoting relationships scale: is this feasible to measure therapeutic alliance when patients see multiple peer workers? Will the instrument focus on the therapeutic alliance with one peer worker of is it a more general measure of alliance?
- Have you considered to measure the working alliance by interviewing the peer workers as well?

As discussed elsewhere, there is only one Peer Support Worker per participant. The Recovery Promoting Relationships Scale (RPRS) comprises two subscales: one measures general therapeutic alliance; the other measures the perceived recovery orientation of the peer worker. A main purpose of using the scale is to explore whether the (participant-rated) recovery orientation of the peer worker is distinct from, or relates to outcomes independently of, therapeutic alliance in general – i.e. whether recovery orientation may be a critical component of the intervention. Assessing the peer worker's perspective of therapeutic alliance, while of interest, was not needed to explore this, and would have added to the overall burden of data collection.

p10.
- Why did you choose a 1:2 allocation rate?
We're not sure where 1: 2 comes from as this is not the correct figure – however, we have improved our explanations under "Sample Size" of our procedures. We in fact used 1:1 allocation in the trial – the way this was arrived at taking into account inflation to allow for clustering in the experimental arm is now explained in the "Sample Size" section (p12).
- Where is the (relatively low) ICC of .03 based on?
This was an estimate provided by the clinical trials unit overseeing the trial, based on similar trials involving clustering by staff member: no ICCs from really similar peer support interventions were available to us. Subsequently a meta-analysis of therapist effects in low intensity interventions has found a pooled ICC around 0.02, indicating that our estimate of 0.03 is conservative . This is noted in the section on Sample Size (p12).

p11.
- changes to the procedures were sufficiently minimal: what were the changes?
Changes were the introduction of additional mechanisms for providing support to Peer Support Workers, and the introduction of measures of loneliness, social outcomes, social capital and social network. No changes of substance were made to the content of the intervention or to trial procedures. We have now noted this in the protocol (p12).

- Have you considered to use stratified randomisation by centre or diagnosis? Why (not)?
We do stratify by site – this is already stated in the section on Group Allocation (p13). Further stratification was not deemed necessary as the sample size is relatively large.

- Will you register characteristics of patients who choose not to participate in the study?

Our ethics approval and local data protection procedures do not allow us to do so.

p14.

- You indicate you will use complete case analysis and refer to this as ITT. However complete case analysis is not ITT. For ITT missing data should be addressed adequately, eg through multiple imputation or through (G)LMM with maximum likelihood estimation, and with variables associated to non-reponse in the model. Please consider adjusting this section accordingly.
We agree with this point. We have now adjusted the General Principles section so that it refers to analysis according to original randomisation rather than to intention to treat (p16).

- is the cluster size of 1 appropriate? shouldn;t it be 1 cluster for all participants?
A single cluster would, we believe, imply some correlation between these patients. We want to assume independence, hence our choice.

- "random effects": do you plan a random slope, random intercept, or both? Please elabore
We agree. Random intercepts are what is meant. This has now been amended in the Primary Outcome and Secondary Outcomes sections (p16).

- Results will be reported as OR while before you mentioned percentages (50% vs 35%). Please consider harmonising this.
This would be difficult to harmonise: the percentages are the basis of the power calculation, but odds ratios are the primary output from the logistic regression selected as most appropriate to test the primary hypothesis.

p15. Linear modeling is not the most appropriate for count data (days spent in care), could you address this?
We agree. We have changed this in the Secondary Outcomes section to Poisson regression with random intercepts (p16).

p16. Please define the acronym PEER SUPPORT WORKER. Is there 1 PEER SUPPORT WORKER per patient or multiple. I had the impression multiple but I may be wrong. If 1 per patient: how are patient and PEER SUPPORT WORKER matched/paired?
PSW means Peer Support Worker – as this is not a very widely familiar abbreviation, we have removed it and substituted the full form throughout. There was only one PEER SUPPORT WORKER per patient. If participants specifically requested a PEER SUPPORT WORKER of their own gender, this is arranged, but no attempt beyond this is made to match PEER SUPPORT WORKERs and participants. There is no consensus in the literature on whether, and on the basis of which characteristics, PEER SUPPORT WORKERs and clients need to be matched. In practice, with three PEER SUPPORT WORKERs available in each CRT, we anticipated being unable to match on many characteristics, and felt that attempting to do so may restrict generalisability to routine NHS settings, where matching is often not feasible. This explanation has been inserted in the section on Delivery of the Intervention (pp 8-9).

p17. Economic evaluation. It is somewhat more common to have the societal perspective analysis as the base care. Please refer to the ISPOR 2015 guidelines or see Drummond et al 2015.

When evaluating new interventions for implementation in the English National Health Service the body responsible for recommending their implementation, the National Institute for Health and Care Excellence (NICE), recommend a health and social care cost perspective over a societal cost perspective. We have made this clearer in the document. We did not collect sufficient data as part of the trial to be able to conduct an analysis from a societal cost perspective that would meet the requirements set out by ISPOR (2010). Instead, aspects of what would be included in a societal

analysis are included in supportive analyses only. We have edited the initial paragraph of the health economics section to clarify our approach (p19).

ISPOR reference: Garrison LP, Mansley EC, Abbott TA, et al. Good research practices for measuring drug costs in cost-effectiveness analyses: a societal perspective: the ISPOR Drug Cost Task Force Report—part II. Value Health 2010;13:8–13)

p18
- first line: 'costed' -> 'valued'
Costed does accurately reflect our processes: We are not assigning a value to the health care resources - we are assigning a cost.

- Confidence intervals: are these 95% intervals?
Yes, they are. This is now clearly specified in the section entitled Confidence Intervals (p20).

- it is somewhat more common to report median ICERS (instead of mean)
Our analysis complies with recommendations set out in Briggs et al (1999) which states that “even when costs are skewed economic analyses should be based on means of distributions”. We are not aware of any guidance documents that recommend the reporting of medians over means for economic evaluations and would question the logic of doing so, especially when adjusting for baseline values.
Briggs AH, Gray AM. Handling uncertainty when performing economic evaluation of healthcare interventions. Health Technol Assess 1999;3(2).

- how will you handle ICERs in case of bootstrapped samples in the dominated quadrant (less effects, more costs)?
In the calculation of the cost-effectiveness acceptability curves results from the bootstrap analysis that fall in the dominated quadrant will be handled the same way as ICERs that are above the willingness to pay for a QALY threshold - they will not be considered cost-effective and hence valued as a 0, where the value of 1 will be allocated to bootstrap results where the ICER is less than the willingness to pay threshold, including dominate results.

- Are other health care contacts (except mental health and GP) also recorded and used in the analysis?
To reduce patient burden responding to questionnaires no other health care contacts were recorded in the completion of questionnaires. We only had access to mental health care medical records, not medical records for other health care contacts. We have made this clearer in the document (p21).

For discussion / general
- It seems that the control group only receives an information booklet while the PEER SUPPORT WORKER group receives care which may be considered 'normal care': making a crisis/emergency plan etc.

Will the control group receive the actual care as ususal in the service area where the study is performed or is the usual care situation more like the care the PEER SUPPORT WORKER group receives?
The care received by the peer support worker group does not reflect ‘normal care’ in the areas: our work in the parallel stream of the CORE programme indicates that structured interventions for self management are not widely implemented. Thus the control and experimental groups both receive something additional to usual care, but the control intervention is much simpler and requires much

less investment of resources than the experimental intervention. This is now noted in the Setting section (pp 5-6).

- CRT workers are blinded to the study condition the patient is in, but what about other (mental) health workers? May this have an impact on your study?
We did not blind CRT workers, but did delay informing them of trial allocation until the point of discharge in order to avoid influencing plans made. PEER SUPPORT WORKERs briefly recorded their visits in notes, so that other involved mental health staff may have been aware of them. Keeping these visits secret was unacceptable to the Trusts for clinical governance reasons, but this does constitute a potential limitation, as awareness of the PEER SUPPORT WORKER could possibly have influenced other care. We will compare care received between arms as part of the health economic analysis. This is now noted in the section on Blinding (p13).

- Please include a discussion/strengths/limitations section near the end of the ms.
Such a section has been included after the abstract, as requested in the editorial comments.

VERSION 2 – REVIEW

REVIEWER	Matthijs Blankers Arkin Mental Health Care - The Netherlands
REVIEW RETURNED	17-Jul-2017

GENERAL COMMENTS	Thank you for addressing our and the other comments. Overall, we are satisfied with the adjustments. We have one final point we would like to draw your attention to: - Although three general points regarding strengths and limitations (actually only 1 limitation is mentioned) have now been included in the manuscript, we think the paper would profit from a somewhat more elaborated discussion of the strengths and foremost the limitations of the study. For example, one could argue that the fact that the peers have not experienced psychiatric crisis themselves means that they are not actual peers who can teach the psychiatric crisis patients new skills and coping strategies acquired in their own recovery journey. This is something we would like to see addressed in this discussion section, together with the other main strengths and limitations (e.g. methodological and practical challenges around perform research among this population and how these are addressed in your study design) the authors experience or expect during this study. For the readers it would be very interesting to hear your thoughts on this.

VERSION 2 – AUTHOR RESPONSE

- We thank Dr Blankijs for his further interesting points. We have extended one of the strengths and added three further limitations, including addressing his comments about the peer support workers' experience of crisis services. We believe that extensive discussion of challenges of acute care is best left to the results paper when we can report our experiences of this, but hope that the paper has been improved by a more extensive Strengths and Limitations list.

- When we resubmitted, I received a notification from the sub-editor that we should make it clear what

out two new supplementary online files are - in accordance with SPIRIT, these are the study consent forms. I have described what they are in the Ethics section at the end of the protocol paper.