**Reviewer’s report**

**Title:** Facilitating Implementation of Research Evidence (FIRE): an international cluster randomised controlled trial to evaluate two models of facilitation informed by the Promoting Action in Research Implementation in Health Services (PARIHS) framework.

**Version:** 1  **Date:** 13 Sep 2018

**Reviewer:** Sarah Cotterill

**Reviewer's report:**

This is an interesting and ambitious trial and we enjoyed reading the manuscript. We were asked to review the statistical analysis, so we have focussed on that aspect of the paper.

Study design summary: this is a cluster randomised controlled trial. 24 clusters (nursing care sites) were randomised to one of three groups (control, dissemination1, dissemination2), stratified by country (x4).

a) The primary outcomes are three measures of 'compliance with continence recommendations' (resident screened; assessed; treatment plan; specialist referral) (abstract and in the methods p11). We are not clear how these are measured. Is this a percentage compliance at the level of the patient or the cluster? (nursing care site). If the latter, the regression models (tables 2 4 6 8) only include 24 units. We have assumed in the following comments that the analysis is based on patient-level outcomes. It would be helpful to expand the description of the outcomes in the methods section to explain how the measures have been constructed. Add a N=   at the bottom of all the tables, to give the overall population, making clear whether these are sites or patients.

b) The analysis of the primary outcomes (three measures of compliance with continence recommendations) uses linear regression models. This is a cluster randomised trial with multiple levels (country, site, staff, patient). In a cluster randomised trial it is usual to perform an analysis which adjusts for clustering at the unit of randomisation, which in this case is nursing care site, and this is what is described. 'Regression' covers a range of different analyses, and in this case presumably some sort of multilevel model was used to adjust the standard errors to take site level clustering into account so it would be helpful to add in exactly which stata command was used to fit the regression models (e.g. mixed, xtmixed). It would be helpful to add a statement that the assumptions of linear regression have been examined and the data meets those assumptions,

c) ICCs were calculated on the baseline data and we find this confusing. Once a multilevel model has been fitted to the data, it is usually possible to extract an overall ICC at the level of
clustering which would be based on all data points (not just baseline) - see https://www.stata.com/features/overview/intraclass-correlations-for-multilevel-models/

d) There were only 24 sites (clusters) and both 'country' and 'intervention' are fitted at a site level and have a total of 5 individual covariates between them - it is doubtful there is enough data here to provide stable estimates: We'd recommend that the authors try removing 'country' from the analysis as a sensitivity analysis to check that results are similar.

e) Patients were recruited to this trial post randomisation - which means that recruiters already knew which arm of the trial a site was in before consenting patients. This lack of allocation concealment can lead to differential recruitment in a cluster randomised trial - e.g. differences in numbers recruited or the type of patients recruited. We can see little information on which to judge whether or not this was a problem. We can't find a consort flow diagram and this would be helpful (and recommended by CONSORT) - how many patients were approached but did not consent, and how many consented but did not provide outcome data? Baseline compliance in the control group was much higher than in the intervention group (table 3) and this could be an indication of differential recruitment. Perhaps the authors can make some comment on this.

f) How similar were patient demographics across groups? The paper would benefit from a table of baseline characteristics, by group. Provide mean(SD) for continuous variables, number(%) for categorical and median (IQR) for ordinal or non-normal, and include both the number included for each measure plus the overall total. It is not good practice to compare the groups using statistical tests (as described on page 11 and shown in table 1).

g) EQ5D-VAS is reported (table 1) at 24 months only (because of least amount of missingness at that time point). It was collected at several time points, suggesting it might be an outcome measure, but it is not reported as such.

h) Table 3 and 5 both suggest an effect in favour of the 2 intervention arms, and the effect is quite close to the 15 percentage point difference anticipated in the power calculation (Table 7 does not). This, combined with the very wide confidence intervals, might indicate that there is a hint of an effect here, but that there is insufficient power to detect a difference. We like the way the authors have used the findings from the process evaluation to explain the results: however, they cannot write off the idea that there may have been an effect, and the lack of power may be more of an issue than suggested in the conclusion.

i) Secondary outcomes at patient level seem to have been reported as a difference between baseline and 24 months, rather than compared by group. These outcomes should be reported
as described in the methods, using Anova or chi squared tests, as appropriate, or even using regression methods

j) Given the lack of power in the study, we would very tentatively suggest that the authors give consideration to combining both arms and comparing facilitation against control, although this analysis was not anticipated and would have to be reported very carefully.

Sarah Cotterill and Sarah Rhodes

**Level of interest**
Please indicate how interesting you found the manuscript:

An article of importance in its field

**Quality of written English**
Please indicate the quality of language in the manuscript:

Acceptable

**Declaration of competing interests**
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

We declare that we have no competing interests.

Sarah Cotterill and Sarah Rhodes
I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal.