Reviewer's report

Title: Training in Dual Diagnosis Interventions (The COMO Study): A Randomised Controlled Trial

Version: 2 Date: 18 June 2007

Reviewer: fiona lobban

Reviewer's report:

General

-----------------------------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

Summary

This is an interesting paper which addresses the important issue of transferring interventions from research settings into the real world of the NHS. However, there are a number of limitations that I feel need to be addressed before it is acceptable for publication.

The intervention
Firstly, the overall aim to train NHS Key Workers in CMHTs to offer a combination of MI and CBT to this very complex client group after 5 days of training, seems extremely optimistic. These are very complex approaches which require high levels of skill and require close expert supervision. More detail of the exact nature of the intervention might help, and on the baseline level of skill of the participants.

Secondly, the case that these approaches have been clearly demonstrated to be effective in a research setting (which would be a prerequisite for dissemination to NHS staff) is not strongly made.

Training
More detail on evaluation of course – if so good – why did people not attend? Those were there at the last session are a very biased sample so cannot conclude it was well received. Need qualitative interviews to follow-up?

Although supervision is mentioned, and the problems of attending are alluded to, more detail is needed to specify how many sessions were attended.

Design
“cluster randomised controlled trial” – what is the unit of cluster? How has this been accounted for in the analyses? It is unclear at the outset, where the unit of randomisation will be the team or the Key Worker. This needs to be clearly stated early on.

The idea that some Key Workers in a team are trained in this approach and others not, seems somewhat odd and it is difficult to imagine how “leaking” would not occur. At the most basic level, although service users have a named care coordinator, they often receive a service from a range to team members to cover shifts, holidays etc. To ensure consistency, it is not clear why whole teams weren’t trained and randomised, as opposed to individual key workers?

In this design, the clustering is alluded to at the level of key worker, however there is no recognition that there will also be a clustering effect at the team level due to variation in the ways teams will support this kind of work.

Power analysis

Sample size calculated on 3 measures
-number of bed days
-whether or not using drugs at follow-up
-Kemp compliance scale
Then has section on main outcome measures – these are staff measures. Not done any power calculation
on the staff measures which is what this paper is about. The fact that power calculations were done for the service user outcomes that are reported elsewhere needs to be made, along with a comment that the study may not therefore be powered to find adequate differences on the key worker measures. Some attempt should be made to assess this for the key worker measures.

The power calculation that is reported concludes that 44 key workers and 220 pts are needed—how? The combined mean of eligible participants according to the consort in both groups is 3—therefore 44 will provide 132 even if everyone took part. Therefore how is the calculation that 44 key workers will result in 220 patients made?

How is the (design effect 1.08 for 5 patients per key workers) calculated? A list of assumptions is given—“pre-post correl of 0.05, intraclass correl between workers of 0.02, and drop out of 20%”—where do all these assumptions come from? Need to justify?

Measures
3 of the 4 main measures were devised for the study with no reliability / validity checks. Internal consistency is reported, but only Cronbach’s alphas given. Mean inter-item correls should be presented (especially as one has alpha of 0.97)

Data analysis
Missing data—at follow-up
-This is very considerable for the intervention and not for the control groups. Implies something about having the intervention meant they did not want to complete data. Suggests is important—therefore how valid to just use reduced sample size? Total n = 63
Why only look at 18mo FU—why not also look at more immediate change—is it not likely that the training had a short term effect that wore off? If this was measured, it needs to be reported, even if the findings were negative. If not, a clear rationale is needed for only assessing at 18 mo FU.

There is no specification of what level of statistical significance is being used, and what corrections have been made for the large number of statistical tests carried out.

Corrections/ expansions

Abstract results
“There were significant differences on 2 of the 6 subscales of the AAPPQ: ‘adequacy of knowledge and skills in working with alcohol and drug users” (adjusted difference 3.598 [95% CI 1.034 to 6.162], f=7.907, p=0.007) and “self-esteem in working with alcohol and drug users” (adjusted difference 3.001 [95% CI 0.460 to 5.542].”

-what direction were these differences in? In favour of experimental group?
-why not reported the f and p value for self esteem as have for knowledge and skills?
-“adjusted difference”—in what way has it been adjusted? For baseline differences?

Last paragraph of background
“The overall aim” sentence reads rather strangely
-what is a worker hypothesis?
-(self…..no end of brackets
Needs proof reading—lot of grammar problems “

Discussion mentions that there was “no demonstrable impact on service user outcomes”—why are these results not reported here? Are they reported elsewhere? Should there be a ref to the Br Jnl Psy paper in press at this point? Could be made clearer in intro that service user outcomes are reported elsewhere and that this focuses on one aspect ie impact of training on key worker outcomes.

I hope these comments prove useful

-----------------------------------------------------------------------------------
Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

-----------------------------------------------------------------------------------
Discretionary Revisions (which the author can choose to ignore)
What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Needs some language corrections before being published

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:
I declare that I have no competing interests