Reviewer’s report

Title: Effectiveness of a tailored implementation strategy to improve adherence to a guideline on mental health problems in occupational health care

Version: 0 Date: 01 Aug 2018

Reviewer: Bryan Rodgers

Reviewer's report:

This paper describes results from analysis of an RCT assessing the implementation of a strategy intended to improve adherence to mental health care guidelines by occupational physicians (OPs) in the Netherlands. The trial included 66 OPs from six regional offices of the occupational health service in South Netherlands. Randomization was at the OP level with approximately equal numbers of intervention and control OPs at each site. The intervention arm OPs received eight two-hour training sessions designed to identify and address barriers to implementation of the mental health guidelines developed by the Netherlands Society of Occupational Physicians. The control arm OPs received no training or 'placebo' intervention. The outcome measures for the present paper were a set of performance indicators (PIs) to assess the care provided by OPs to individual workers. The analyses reported in the present paper used data from 114 workers who had taken sick leave and agreed to participate (56 intervention and 58 control). They received care from 34 of the study OPs (16 intervention and 18 control). The medical records of the workers were evaluated for all consultations with their OP over the period of sick leave up to one year. Twelve PIs falling into five "key" domains were rated independently by two assessors. Statistical analyses followed the RCT design of the study to contrast each indicator between the intervention and control arms, with consideration given to the clustering of OPs. The main findings were that the intervention OPs showed significantly greater adherence to the guidelines on six of the twelve PIs. Significant differences were also found for four of the five domains of PIs: Process diagnosis, Problem orientation, Interventions/treatment and Relapse prevention, but not for Continuity of care. Overall, however, adherence to the guidelines was low in both groups with adequate adherence falling as low as 2% on one of the PIs and below 10% on half of the twelve PIs. The authors conclude that although the implementation improved adherence, the "adherence to the guideline recommendations is still far from optimal and needs to improve to guarantee high quality of occupational mental health care".

The strengths of the present study include the RCT design, the thoroughness of the implementation strategy (i.e. training), the outcome measures (PIs) developed and utilised which intentionally map on to the guidelines, and the attention in data analysis to possible clustering effects of OPs (which turned out not to be an issue). Weaknesses include the achieved sample size, evidenced by wide confidence intervals, and the possibility of contamination from physicians in the intervention arm to their control colleagues (as acknowledged in the Discussion). It is a trade-off between the latter and the risk of having different types of workers under different area offices and the authors have likely made the better call on that decision. Another acknowledged weakness is that OPs' records may not reflect accurately their findings and activities. A further possible weakness (I could not discern from all the references to
"clustering" whether this has been taken into account or not) is that the OPs are the real unit of analysis in the present study, not the workers. There are data from 114 workers but only 34 OPs. Did the statistical analyses take account of the nesting of workers within OPs? This is a much greater threat to the design that the clustering of OPs in their office settings. (My apologies if I have misunderstood this and that the former is the type of clustering already taken into account. If so, please clarify this in the description of the statistical analyses.)

There are some difficulties with the scope of the present manuscript and the way it is presented. Some of these are serious issues and affect the core of the interpretation of the findings and the suggestions made for the future of implementation of guidelines and for research on this. I note from the declarations of interests that two of the authors were involved in the development of the present guidelines but the level of their involvement is not stated (other than not receiving payment from their subsequent use).

The main problem I encountered with the present manuscript centres on the separation of the current paper from previously reported findings from the same study. (I note that the authors express this as "alongside a cluster RCT" but there is only one study with one registered protocol.) This gives an artificial feel to the paper (the elephant in the room is not unacknowledged in any way). Further, after reading some of the previous publications and the study protocol, I found it very hard to understand the interpretation of results in the manuscript. First, the trial did not achieve expected outcomes in terms of workers' wellbeing (primarily return to work). The publication detailing this is in the references of the present paper (#39) but is not referred to anywhere in the text and the main results are not discernible from the manuscript. Of course, the aims of the present paper are not in regard to the workers' outcomes but it is very strange that this does not get a mention given the content of the Introduction and Discussion sections. A second publication published in 2017 (#42) is more problematic. The same group of authors reported that, irrespective of the intervention, adherence to the guidelines by OPs is not associated with workers' outcomes. This being the case, what possible value is there in adherence or, indeed, strategies to encourage adherence? The manuscript is written from the perspective that adherence to guidelines is necessarily a good thing, e.g. see first paragraph of the Introduction, and the final sentence of the Abstract states "adherence to the guideline recommendations is still far from optimal and needs to improve to guarantee high quality of occupational mental health care.") This presumption of guideline adherence clearly does not apply to the current circumstances. It is possible that the vague final sentence of the Conclusions section of the present paper is alluding to this point, but, if so, it will be lost on the large majority of readers if they are not provided the information to draw their own conclusions. Are the authors in fact concluding that the NVAB guidelines have little value? This IS highly relevant to the particular aims of the present paper because it is understandable that OPs would not adhere to guidelines if they themselves perceive them to be irrelevant to workers' outcomes. Even if this is not actually the view of OPs (the authors would know better than I) it is still a necessary issue to acknowledge and discuss, as it is one plausible explanation for the study's findings (across all of the papers).

Something else that is necessary for readers outside of, or unfamiliar with, the OP system in the Netherlands is a little more explanation of OPs' role in relation to other health care providers. To what extent do they provide mental health treatment themselves or facilitate access to mental
health care from other professionals for the workers affected? I could not get a good feel for this and trying to read between the lines was as confusing as helpful. I note for example, that the PIs relating to continuity of care (5.2) include "Consultations with the worker take place every 3 weeks during the first three months of sickness absence. Thereafter consultations take place every 6 weeks." For workers who have a mental health problem serious enough to prevent them from going to work, this seems highly inadequate if it is the only source of support. Can I assume that they must actually be receiving other treatment (from GPs or specialists)? Why are there no guidelines/PIs on how frequently these workers receive evidenced based good quality mental health care? Playing devil's advocate, PI 3.2 seems the only PI out of 12 that I might anticipate to have some benefit for workers' wellbeing/recovery. No amount of monitoring and recording will improve outcomes in the absence of effective treatments and/or self-help strategies (neither of which appear to be monitored).

As well as the main points above there are a few minor issues that the authors could take on board if they revise the paper.

1) There are a few places where English language does not quite work. In the abstract "little is known" in place of "few is known". Generally, though, the paper is extremely well written.

2) The terminology of "key PIs" is a little confusing as this suggests that some of the 12 PIs are more important than others. Here the PIs are grouped into five domains (or areas) and it would be better to use a term indicating a higher-level grouping rather than "key".

3) Participant workers' inclusion criteria are provided on page 9. I saw from the registered protocol on line that three exclusion criteria were applied to the present study but could not see these in the present manuscript. Were they implemented? If so, they should be included in the paper. Further, how are these exclusion criteria justified, as censoring on the main outcome variable would work against the demonstration of differences between the intervention and control arms of the trial:
   2. Sick leave duration of less than 4 weeks,
   3. Sick leave duration of more than 8 weeks.

4) On page 11, the section covering statistical analysis says "The intention-to-treat principle was used in the analyses." I am puzzled as to how this would be operationalized for the analyses reported in the present paper. No workers' outcomes are analyzed so their data are irrelevant here. As to the OPs' the OPs who did not see any participating workers could not provide data from records, so only those who did see participating workers (n=34) could be included in the analyses. They can only be assigned to the intervention and control arms of the trial in accordance to their original allocation in the design (as there is no other categorization to use), so what does "intention to treat" signify?

5) Some figures in Table 5 are clearly incorrect, particularly risk differences and their confidence intervals. I think this is due to the inconsistent use of minus signs for those differences where the % is higher for the control arm. There are also some instances where a the decimal place has been omitted. In the bottom right-hand cell, I do not understand how the CI can be [0,0].
6) The consideration of possible "barriers" in the Discussion is not always clear to make the distinction between barriers to guideline adherence and the constraints (particularly external constraints) which nullify the potential benefits of guideline adherence. These are two conceptually different things and would show up in different ways in the design of the current study. There is a further lack of specificity in the Discussion where the word "might" is often used in contexts where the present study has not only collected data but already published findings in previous journal articles. It seems odd to speculate about things already reported on. Overall, I can see value in reporting (especially from an RCT) how a systematic intervention does or does not change the behaviour of professionals and why it can work and why it might not work as well as expected. However, it is very confusing to find these results presented in isolation from the context of the full study and, in some instances, with implications and conclusions that conflict with earlier publications from the team.

**Are the methods appropriate and well described?**
If not, please specify what is required in your comments to the authors.

No

**Does the work include the necessary controls?**
If not, please specify which controls are required in your comments to the authors.

Yes

**Are the conclusions drawn adequately supported by the data shown?**
If not, please explain in your comments to the authors.

No

**Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?**
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

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

Needs some language corrections before being published

**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.

I declare that I have no competing interests.

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

Do you want to get recognition for reviewing this manuscript? Add a record of this review to Publons to track and showcase your reviewing expertise across the world’s journals. Signing up is quick, easy and free!

No