Reviewer's report

Title: Effectiveness of improving working conditions on employees' health: A systematic review of organizational-level workplace interventions

Version: 1 Date: 25 November 2013

Reviewer: Johan Abildgaard

Reviewer's report:

Major Compulsory Revisions:
As the study covers vast area of research and different fields of research substantially increased clarity and more thorough definitions are necessary to make the paper more accessible for readers and potentially allow linking the results to one's own research, especially the following:

1. The paper is a review of the evidence of organizational level interventions to improve the working conditions and employees’ health. Though relevant there are a number of elements in the paper that need to be clarified.

The stated objectives are:

(1) to increase the comparability of results of interventions with current labor market-related surveys by making use of harmonized classification schemes,
(2) to evaluate the effectiveness of organizational level interventions with respect to working conditions improvements and employee health, and
(3) to evaluate how the changes in working conditions and the occupational class of employees might be associated with the probability of reporting statistically significant intervention effects

The main problem of the paper is that the link between these three stated objectives and the discussion/conclusion should be clearer. The abstract states that “this systematic review assesses which working conditions facilitate improvements of employees’ health in organizational-level interventions” a very ambitions task, and the findings of which working conditions are to be targeted, should be followed up more clearly in the discussion and conclusion. The conclusion likewise raises questions about “systematic risk assessment in workplace interventions” (page 13) and suggests that organizations “prioritize holistic intervention approaches that identify systematically the specific health risks within and across occupational classes”, not that I disagree but these recommendations seem too far from the purpose of the paper. Perhaps even the results in the paper are better presented in separate papers, as the present focus and research questions are quite ambitions for one 14 page manuscript.

2. Another central issue is the clarity of the studied intervention outcomes. The statement on page 3 of focusing on “organization level interventions and health related outcomes” is imprecise and could be specified more thoroughly,
especially as the definitions of each of these factors could be more specific. Injury prevention (ref 16) is listed which suggests that your definition of what is health related is very broad. Likewise is the reference of a study focusing on “psychosocial and health effects” (ref 17) somewhat vague and should be made more specific. On page 10 health related outcomes is mentioned again and range from perceived health, absence and heart disease. The paper would improve substantially with a clearer definition of “health related effects”.

3. The selection criteria (2) on page 5 state that only studies reporting health related variables are included. Does this cover all the variables listed on page 3, if so I would assume a more precise definition and clarification would be needed of, for instance “psychosocial and health effects”. Likewise a clear definition of (page 5) (1) “strong organizational approach” would be an improvement.

4. Similarly the section on page 3 on comparability of interventions could be improved by addressing how some researchers problematize relying too heavily on standardization of interventions and instead argue for concepts such as intervention fit (Nielsen & Randall, 2012). For instance, research question 1 argues for “making use of harmonized classification schemes”, which might be a good idea but in order to assess the potential use of this technique in other research more information on how this is done is needed. I agree that more specific reporting of population job types, and perhaps some agreed upon common measures of health would be preferential but if this is the key argument of the paper it should be unfolded more and argued more clearly.

5. Considerations about the quality of data and the problems of evaluating interventions on a large scale company level with RCT designs (a method originating from individual treatment in medical science) is not addressed. At least the motivation for a priori deeming all quasi experiments (Which have been praised in (Grant & Wall, 2009)) and other non RCT studies to be of low quality should be clarified. Likewise the criteria for when a RCT design is achieved should be made clear. E.g. is randomization between 2 departments in one company RCT? Preferably, but not necessarily, the discussions on the suitability of RCT design in conducting organizational level interventions could be noted.

6. Though it is mentioned that some of the categories on page 6 correlate, the solution of rotating the factors seems to avoid addressing the potential that some of these factors are not just correlated but different sides of the same aspect. Time related conditions (task factors) psychosocial conditions (perceived conditions) and work organization (structural factors) might in some cases (arguably) be the same features conceptualized on different ontological levels and therefore should be highly correlated. The definitions of the categories and the motivation for reducing correlation through striving for orthogonality should be substantiated.

Minor Essential Revisions

1. Regarding the weighing criteria of intervention quality it could be noted that sample size, response rate and other quantitative criteria used in the present paper are not the only quality criteria used in intervention research. Several researchers and reviews have called for more qualitative evaluation and
assessment of the context and process of implementation (Egan, Bambra, Petticrew, & Whitehead, 2009; Nielsen & Abildgaard, 2013; Nielsen, 2013). The presented quality criteria might be more appropriately stated as being quantitative evaluation criteria or quality criteria for quantitative data and not quality of the studies per se. Likewise the weighing of high N studies should be argued more clearly as it assumes that the quality of assessment is similar to the smaller studies and that they are comparable in their approach and evaluation methodology. One might suspect that studies with for instance 3000+ employees will use more general interventions and assessments than studies with less than 100. These differences are not considered and the high N studies are weighed as more substantial as would be valid if it was studies of individual interventions.

2. The result that the middle and lower quality studies have a higher amount of significant effects than the high quality studies which is puzzling. This finding should be discussed in the paper.

3. The potential of circular results regarding significant effects and number of changed factors is not considered. One might suspect that some studies might omit factors that were initially thought to be targets but failed to improve due to either theory or program failure. There are general discussions in the intervention field regarding the publication of null and negative effects being under-published (see for instance Olsen et al., 2008). If under-publication of negative findings is taking place the result that number of factors reported correlates with number of effects changed is not surprising. Considerations about reporting of results in the included studies would thus improve the paper.

Discretionary Revisions

1. The studies [ref 9-12] on page 3 could preferably be supplemented with newer literature. Either newer intervention studies or preferably some of the more recent broader analyses of the intervention literature (Barends, Janssen, Have, & Have, 2013; Nielsen, 2013; Semmer, 2011).


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

Quality of written English: Acceptable

Statistical review: Yes, but I do not feel adequately qualified to assess the statistics.

Declaration of competing interests:

I declare that I have no competing interests