Reviewer’s report

Title: Comparison of alternative versions of the job demand-control scales in 17 European cohort studies: the IPD-Work consortium

Version: 1 Date: 2 September 2011

Reviewer: Rosane Harter Griep

Reviewer’s report:

1. Is the question posed by the authors well defined? Yes
2. Are the methods appropriate and well described? No
3. Are the data sound? No
4. Does the manuscript adhere to the relevant standards for reporting and data deposition? No
5. Are the discussion and conclusions well balanced and adequately supported by the data? No
6. Are limitations of the work clearly stated? Yes
7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished? Yes
8. Do the title and abstract accurately convey what has been found? Yes
9. Is the writing acceptable? Yes

Congratulations on the authors’ initiative to build the IPD-Work Consortium. It is certainly a study with so many possibilities to elucidate associations which are still unclear about the influence on health effects of psychosocial risk factors at work.

1. Major Compulsory Revisions:

Although the article is clear in its purpose and relevance regarding to the main question of the study, I would like to mention some concerns and to request some clarification in relation to this article:

In statistical analysis proposals, the authors use Pearson correlation coefficient and its confidence intervals. In the context of this study, this indicator means that part of the scales are correlated with a whole, also it requires just a linear relation between the two measures. I’m not sure that this information answers or helps to answer the main question of the study: the use JCQ or DCS instruments in full or partial scales guarantees that we are using instruments with similar psychometric properties, comparable and therefore all forms are capturing the same construct.

As the use of indicators refers sensitivity, specificity and kappa statistics, the authors propose to verify the validity with respect to only one form of use of the scale, i.e. combination of quadrants. However, there are assessment and other
uses, widespread in the literature, which also interferes in the comparability between studies, such as: demands/control ratio, strain by subtraction, quotient approach, evaluation of each dimension (demands/control) in separated, distribution in quartiles or tertile. (Please read: Courvoisier & Perneger, 2010; Landsbergis et al., 1994).

Actually there is no agreement regarding to the best approach to assess job strain based on the demand control model, according to Karasek’s model (“quadrant” term). Other formulations have been introduced, especially the ratio between demand and control to define stress at work. This “quotient” term generates a continuous form of the variable that can be dichotomized at an arbitrary cut-off point (Landsbergis et al, 1994; Courvoisier & Perneger, 2010). In fact, some authors have applied this approach using different cut-points, including the median value (Goldstein et al, 1999; Clays et al, 2007), the third tertile (Tsutsumi et al, 2001) or the forth quartile of the distribution (Theorell et al, 1988). Despite the large number of publications on this subject, there is no agreement regarding the best procedure to evaluate job strain.

Given the importance of the topic of this article and in future publications of the group on the topic, I suggest more in-depth evaluations of the psychometric characteristics of the several formulations of instruments presented and clearer about the implications influence of the choice of instrument (JCQ x DCM) as well as partial version in their health outcomes.

Please, explain how the translation and cross-cultural adaptation of instruments used in each country were done.

Clarify the reasons of the use of different instruments across countries (studies) that participate in the IPD-work Consortium and if there was any reason or criteria in choosing items of the instrument in each study.

Clarify what is written in "definition of the complete and partial scales of job demands and control". "We constructed partial scales using JCQ/DCM items (or item judged to be very similar to these) that were available in each of the IPD work studies that did not have the complete scales". How was this judgment done? Who judged it?

Table 2: the authors mentioned “other” scale (DWECS, NWCS, Pols, PUMA, and Still Working). There is no comment in text about what “other” instrument means. Please, clarify.

Include in table 1 the items exactly the way as they had been collected (statement/questions and answer patterns). Thus, it would become clear the differences to the readers: “the JCQ items are expressed as statements and the respondents are asked to report if they agree or disagree with the statement on a four level Likert scale, while DCQ items are expressed as questions with the response scale being frequency based”

References:

Goldstein I, Shapiro D, Chicz-DeMet A, Guthrie D. Ambulatory blood pressure, heart rate, and neuroendocrine responses in women nurses during work and off
work days. Psychosomatic Medicine. 1999; 61: 387- 396;


**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.