Reviewer's report

Title: Assessment of long-term sickness absence: content and face validity of a new questionnaire based on qualitative data from nominal groups

Version: 1 Date: 30 Apr 2019

Reviewer: Katherine Jones

Reviewer's report:

General Comment:

This paper has been improved from the first review. However, I believe it requires additional improvement prior to publication. Specifically, the authors' interchangeable use of the terms item, factor, and category may have led them to overlook overarching patterns that emerge from their qualitative methods. Specifically, they conclude on p. 16, line 359, "Further research should decide which factors are effective in predicting long-term sickness absence." However, in their background, p. 2, lines 52-54, they state that these factors are already known. I believe their findings indicate that future research should test whether adaptive work environments that address the patient's perceptions of their illness and health can minimize long-term sickness absence.

Specific Comments:

First, the fluency and understanding of the paper is much improved after revision. However, some awkward constructions and confusing use of terminology remain.

* For example, p. 7, lines 150-151, "Each participant could read only one post-it at the time to avoid that some participants would not get the chance to speak." May be improved to "Each participant could read only one post-it at a time to ensure that all participants had the chance to speak." On p. 14, line 300 the authors describe the legislative environment of the patient, which is very confusing. Consider the following: "Next, we noted that two extra items concerned the patient's knowledge of regulations governing the return to work process."

* The authors appear to use the terms "factor," "item," and "reason" interchangeably at times, which leads to confusion regarding the 61 items in the QuickscanENG and which factors that may be predictive of long-term sickness absence they measure (load upon). Specifically, on p. 2, the authors state, "Several literature reviews have identified factors predicting long-term sickness absence, such as gender, age, level of education, marital status, number of children and private life strains, perceived health, mental and psychosomatic complaints." These are the
characteristics or factors associated with long-term sickness absence. However, the authors also state that the 61 items of the QuickscanENG load on 21 predictors. It may be best to use the term factor rather than predictor here (p. 2, lines 62-63). However, in the abstract and discussion, the authors refer to items when they may mean factors. Specifically, Abstract p. 1, lines 25-26, "Qualitative data were collected with the nominal group technique. The data were gathered exploring items that influence return-to-work restrictions or opportunities." And in the title of Table 2, p. 11, "Codebook of items;" and in the discussion, p. 13, lines 256-258, "Hence, 20 out of 21 items were spontaneously discussed by the participants as important factors, which might encourage or prevent them from resuming work." This is confusing because the QuickscanENG has 61 items that load on 21 factors. In Table 3, the authors indicate that 61 items are grouped into four categories (Work-related factors, Stressful life-events, Functioning, and Person-related factors). And on p. 12, lines 240-241, the authors state, "Twenty-six factors out of 37 were selected in the ranking as the most important factors across all participants, indicating the diversity of priorities for participants."

I recommend the following:

1. When referring to the "quickscan" instrument, use the same terminology throughout; preferably "QuickscanENG" as in the supplemental material. Make it clear in the Abstract and background that you (the authors) developed the QuickscanENG from a review of the literature and existing instruments as described in Table 3, supplementary material.

2. Organize the 61 items in the Supplementary QuickscanENG into their 21 factors and only use the term "factor" to refer to these 21 factors.

3. Provide a definition of each of the 21 factors and four categories rather than just an example of an item that loads on a factor.

4. Explain why you grouped specific "factors" within one of the four over-arching "categories."

5. Upload the structured questions used to facilitate the focus groups as supplemental material so that the reader can judge the language used to elicit participant responses. Specifically, do the authors ask about "factors" or "reasons" for returning/not returning to work. It would seem that the focus group generated reasons, which then were categorized into the various a priori factors for subsequent validation via factor analysis. Use the term "reason(s)" or some other noun that the participants used when describing what participants of the focus groups prioritize.

Regarding tables and figures, the new column headings in Table 2 are helpful but it is confusing to use the column heading "Question" and the term "item" in the Table title. It does not seem that
these questions in Table 2 match the items in the QuickscanENG so this is very confusing to me. Figure 1 is now legible, but to what do the values in each piece of pie refer? Improve the title of the figure to be more descriptive (e.g. Top 10 Priorities of Focus Group Participants to Facilitate Return to Work). Recall that every table and figure should stand alone.

Second, the purpose statement is improved. However, I recommend removing the reference to the literature, on p. 2, line 54 so that the sentence reads, "A limitation of existing questionnaires is their focus on a specific medical condition, despite experts agreeing that specific medical conditions are not the most important predictor of long-term sickness absence." I also recommend clarifying whether it was the authors who developed the QuickscanENG. This clarification can be provided by using active as opposed to passive voice ...."Therefore, based on these predictors identified in the literature and existing validated questionnaires, we developed and validated (factorial validity, concurrent validity) a new questionnaire called the QuickscanENG to assess the risk of long-term sickness absence in Belgium [8]."

Third, the authors have clarified that they used an integrated approach to coding and they have included some quotes from the focus group participants. However, without specific definitions of the factors/coding themes, it is difficult for a reader to judge whether a quote is representative of a specific factor/coding theme. When reporting results, it would be helpful to better organize the text. For example, on p. 8, lines 172-174 could be reworded as: Six factors were discussed in all five focus groups. Three of these factors were work-related (social support by management, social support by colleagues, and learning and development opportunities) and three were related to functioning (psychological distress, pain perception, and patient's health perception). Then provide representative example of the quotes in the same order as they are listed in the text. Use this same pattern as parallel construction to report all of the results. It would be most helpful if quotes that are representative of the 20 factors in the QuickscanENG that were identified in the focus groups were reported in a table labeled with those factors. Finally, clarify that four categories are reported in supplementary Table 3 but that all codes were assigned to five categories because the category "Environmental Factors" emerged during constant comparative analysis (see p. 8, lines 169-171). If the authors report their results by consistently identifying how quotes related to codes/themes/factors can be aggregated into five overarching categories then patterns emerge; each patient has a unique situation but it is likely that there are overarching priorities as revealed by Figure 1. Specifically, work-related factors and the patient's functioning are common priorities. Finally, given that the authors did not use self-determination theory to develop the QuickscanENG or to code the focus group textual data, I believe it may be better to identify this relationship as a focus of future research. For example, ask the research question: are autonomy, relatedness, and competence maximized for those returning to work if
management and colleagues are supportive and the work-environment is adaptive (i.e. provides opportunities for learning and development)?

Fourth, I did not mean to imply in my first review that a theory is not the final result of a study using an integrated approach to coding as the authors now explain. If the authors more consistently relate the themes to the 20 factors in the QuickscanENG and then to the five overarching categories, I believe a theory ("if, then" format) appropriate for testing using quantitative methods may emerge. For example, "if work-related factors such as an adaptive work environment and support from management address how the patient perceives his/her illness and health, then long-term sickness absence is likely to be minimized." This theory recognizes that the top reasons for long term sickness absence as reported in Figure 1 are categorized as work-related and functioning (illness recognition and patient health perception).

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

Yes

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

Unable to assess

**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.

Not relevant to this manuscript

**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!