Reviewer’s report

Title: Methods for developing a theory informed patient reported outcome measure: the example of a measure for glaucoma screening

Version: 1 Date: 25 February 2011

Reviewer: Benoit Arnould

Reviewer’s report:

This is a good manuscript, quite long but easy to read, matching the current standards, worth the publication, contributing to addressing the current challenge of creating new measurement instruments in the field of ophthalmology. The methods are well described and consistent, the work is based on a reasonable set of data, clearly documented.

However, there are a couple of important points that, in the reviewer’s opinion, call for essential revisions:

The first one is about the title, as well as about the abstract introduction and conclusions: the title suggests this is a paper discussing the methods, illustrated by a specific measure in glaucoma. I think the authors should give a clear priority to their paper: if they advocate for a specific method in developing new PRO measurement instruments, the discussion on the theories (those that have been used in the past, those currently used for new instruments, and the specific approach they suggest) should be more developed and supported by more references and an accurate discussion, which is not the case here; if they essentially promote a new instrument – which seems to be the case – then the title is inadequate, as are the statements in the abstract introduction “we describe generalisable method…” and conclusions “the study addresses a gap in the literature”.

The second major limitation in the work – which is related to the one mentioned above – is about the term “theory”. In qualitative research, a “theory” is frequently referring to “grounded theory”, which claims that the theoretical conceptual framework should emerge from data coming directly and essentially from the information coming from the study subjects. This is not the case here, as the theory is pre-existing: the authors use the ICF theory as the basis for the interpretation and organization of items, rather than a theory which would have emerge from specific qualitative research with patients. We found no justification in the manuscript explaining why the authors did consider a preexisting model would be superior to a grounded theory model. Moreover, the validity of the preexisting model itself is in our opinion not well justified: why this model rather than another one; how was this model developed, why is it a good start for developing a new scale?

The third major limitation is about the use of this new instruments: “screening”
suggests this instrument is aimed at detecting specific patients among a more general population, who would be eligible for a specific action or decision. What is the clinical or epidemiology question that the new questionnaire is intended to address? Is it treatment, is it referral, is it communication, or anything else? If there is an intention to screen for specific patients, what is the next step: threshold, decision rule, classification… And, what is the targeted population: general population, population already showing some risk factors, population complaining about the something, population visiting specialists?

The fourth limitation is in the selection of information and questionnaires used for the development of the current scale: one of the selection criteria is availability of the source questionnaires in English. How can the authors claim this is generalisable, if they did not assess the contents of other existing questionnaires developed with non English-speaking patients?

In addition, the following discretionary revisions are suggested:

Page 4: “the optimal glaucoma-specific PRO instrument evaluating a screening programme would comprise items relevant to disability associated with glaucoma and its treatment and items capturing the impact of any disability on quality of life”. Why so? How do authors justify this statement?

Table 1: Step 3D: how do you justify that items are not relevant to the aim of the instrument?

Step 3 in text (page 8): item reduction. Step D: not really informative. Step E: how did the authors ensure reproductibility?

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

**Quality of written English:** Acceptable

**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 interest