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: 22 February 2011

Reviewer: Ecosse Lamoureux

Reviewer's report:

I have serious reservations about the use of the term ‘item bank’ throughout this paper. It is my belief that the authors here refer to a pool of items that they have amassed throughout literature and other means. An item bank is the end product of several important developmental and validation phases where resulting items have appropriate fit; have reasonable targeting; are well calibrated; and have undergone some form of item response theory. What the authors claim to be an item bank in the paper is simply an initial pool of items. It shows a serious misunderstanding of the appropriate terminology.

I also surprised that the focus of the AGQ is very narrow as the authors are only interested in functioning and disability which are similar. To generate an item bank with such a narrow focus does not justify the invested resources and time. What are the authors’ plan following this?

Introduction

I fail to make the connection between glaucoma screening and a valid glaucoma instrument. It is one of the reasons but there are dozens more. One line is enough and other reasons also need to be mentioned. In addition the authors claim further that they want the tool to be used in clinical trials. Are the items geared towards screening or treatment and clinical trials?

Methods

The authors report ‘From the resultant list of instruments, we selected those instruments that met the following criteria: suitable for self report; validated in a glaucoma population; in the public domain; items and response options fully described in the text article reporting the instrument’. Rather than using the term ‘instruments’, it should be items. The authors are seeking to generate a pool of items, as such they should focus on items not instruments.

The authors argue the following: C) It is recommended that both generic and condition-specific PRO instruments be used in clinical trials [26]. The glaucoma-specific AGQ was therefore designed to be administered alongside the widely validated generic measure, SF-36, in the definitive trial [26]. The third phase of de-duplication resulted from this decision. All SF-36 items and any extra items considered, by the multi-disciplinary team, to be directly covered by SF-36 items were removed from the item bank.
What is the rationale for including generic instrument? Just because they have been recommended? How does including a generic instrument contribute to the aim of the study which is ‘(to develop a self-report measure of functioning and disability associated with glaucoma, and its treatment, for use in a clinical trial)’.

I have serious reservations concerning the two strategies used by the authors to item reduce the pool. It is too ‘expert’ driven as opposed to patient driven. How about pilot testing these items with a group with glaucoma, look for inter-item correlation, poorly understood items, those with little variance, those with high floor and ceiling effects, etc.

The ‘think aloud section’ is encouraging as it is the only patient involvement in this preliminary phase but it was limited to only a group of eight subjects. Clearly more is required.

Result
I am staggered by the final figure of 68 items from an initial pool of 725.

Discussion
One of the limitations of the AGQ, even at this early stage, is its relatively low overall number of items in the pool. This will invariably shrink in the future developmental and validation phases. Yet, this is supposed to be an item bank with a range of items able to address functioning and disability across the spectrum of the condition. I think the authors have missed a tremendous opportunity here. I suspect not having done focus groups from the start, was always going to be an issue in the end.

How do the authors plan the next stages?

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

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:
i declare that I have no competing interests