Reviewer's report

Title: Development and Evaluation of a Quality Score for Abstracts

Authors:
Antje Timmer (antje.timmer@klinik.uni-regensburg.de)
Lloyd R Sutherland (lsutherl@acs.ucalgary.ca)
Robert J Hillsden (rhillsden@acs.ucalgary.ca)

Version: 2 Date: 13 Sep 2002

Reviewer: Dr Roberta Scherer

Level of interest: A paper whose findings are important to those with closely related research interests

Advice on publication: Other (see below)

In this paper the authors present an instrument developed to assess the quality of abstracts submitted for presentation at a gastroenterologic scientific meeting. The instrument designed comprises 19 items and an additional component to balance scores across abstracts describing different types of scientific study designs. To my knowledge, this is the first attempt to develop an instrument specifically designed to evaluate abstract quality of abstracts written for a meeting presentation. As such, it is an important first step.

Compulsory Revisions

1. The authors used reasonable resources in development of the instrument, but most of the resources used evaluated clinical research and usually controlled clinical trials. The final instrument reflects this bias in that the instrument appears to have good psychometrics in evaluating the quality of abstracts describing clinical research, but sub-optimal ones in evaluating the quality of abstracts describing basic science. This dichotomy is present in the inter-rater reliability, the sensibility scoring, and in the raw score without the "add-on" item giving additional points for clinical research design types.

No justification is given for having a single instrument to judge both clinical and basic science abstracts. It seems optimistic to expect to be able to evaluate quality for both types of research with a single instrument, and the dichotomy in the psychometrics argues that it was not successfully done. While scientific quality can be defined by some very basic scientific principles (e.g., items leading to internal validity), there are some quality items specific to clinical research that cannot be applied to basic research - and the authors have acknowledged these. Are there not also some items that can be applied to basic science that may not apply to clinical research, and should be included in an instrument assessing the quality of abstract describing basic science research studies? For example, "is the model used to study a biological system adequate to answer the question?" is not quite the same question as "design evident and appropriate to answer the question?" and might be an important attribute quality for a basic science research study, but not apply to clinical research.

2. Methods for construct validity. The authors are using a circular argument is that they state that they hypothesize that "higher abstract scores should correlate with abstract acceptance" but then they use the fact that an abstract was accepted to evaluate the construct validity of the items used to
determine quality.

3. This instrument was designed to evaluate quality of abstracts written for this specific gastroenterology meeting. Without testing it on abstracts submitted to other specialty or general medical meetings, it is not appropriate to imply that this instrument would be valid across all specialities

**Competing interests:**

None declared.