ELSEVIER

Journal of Clinical Epidemiology ■ (2015) ■

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

An emotional functioning item bank of 24 items for computerized adaptive testing (CAT) was established

Morten Aa. Petersen[a,*], Eva-Maria Gamper[b], Anna Costantini[c], Johannes M. Giesinger[b], Bernhard Holzner[b], Colin Johnson[d], Monika Sztankay[b], Teresa Young[e], Mogens Groenvold[a,f], on behalf of the EORTC Quality of Life Group

[a]*The Research Unit, Department of Palliative Medicine, Bispebjerg Hospital, University of Copenhagen, Bispebjerg Bakke 23, 2400 Copenhagen, Denmark*
[b]*Department of Psychiatry, Psychotherapy and Psychosomatics, Medical University of Innsbruck, Christoph-Probst-Platz 1, Innrain 52, A - 6020 Innsbruck, Austria*
[c]*Psycho-oncology Unit, Sant'Andrea Hospital, Department of Oncological Sciences, Faculty of Medicine and Psychology, Sapienza University, Via di Grottarossa 1035 - 00189 Rome, Italy*
[d]*Surgical Unit, University of Southampton, University Road, Highfield, Southampton SO17 1BJ, UK*
[e]*Lynda Jackson Macmillan Centre, Mount Vernon Cancer Centre, Northwood, Middlesex HA6 2RN, UK*
[f]*Institute of Public Health, University of Copenhagen, Øster Farimagsgade 5, P.O. Box 2099, 1014 Copenhagen, Denmark*

Accepted 7 September 2015; Published online xxxx

Abstract

Objective: To improve measurement precision, the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Group is developing an item bank for computerized adaptive testing (CAT) of emotional functioning (EF). The item bank will be within the conceptual framework of the widely used EORTC Quality of Life questionnaire (QLQ-C30).

Study Design and Setting: On the basis of literature search and evaluations by international samples of experts and cancer patients, 38 candidate items were developed. The psychometric properties of the items were evaluated in a large international sample of cancer patients. This included evaluations of dimensionality, item response theory (IRT) model fit, differential item functioning (DIF), and of measurement precision/statistical power.

Results: Responses were obtained from 1,023 cancer patients from four countries. The evaluations showed that 24 items could be included in a unidimensional IRT model. DIF did not seem to have any significant impact on the estimation of EF. Evaluations indicated that the CAT measure may reduce sample size requirements by up to 50% compared to the QLQ-C30 EF scale without reducing power.

Conclusion: On the basis of thorough psychometric evaluations, we have established an EF item bank of 24 items. This will allow for more precise and flexible measurement of EF, while maintaining backward compatibility with the QLQ-C30 EF scale.

Keywords: Computerized adaptive testing; EORTC QLQ-C30; Emotional functioning; Item response theory; Oncology; Patient-reported outcome

1. Introduction

Computerized adaptive testing (CAT) is a form of intelligent questionnaire; the basic idea is to maximize the precision by only asking questions relevant for the individual [1−3]. For example, if a patient has reported severe emotional problems to the previous items (questions), the next item will be one relevant for patients with severe problems. In this sense, the questionnaire is adapted ''on-the-fly'' to the individual using previous responses to select the most informative next item. Clearly, such adaptation cannot be done using usual paper questionnaires but requires the use of computer technology. All items used in a CAT are selected from a collection of items called an item bank or item pool. In a CAT item bank, the items have been calibrated (fitted) to an item response theory (IRT) model [4,5]. This means that scores based on any subset of the items are comparable. This unique property facilitates the adaptation to the individual without compromising comparability across individuals. The adaptability, that is, the selection of the most informative item at each step, generally makes CAT instruments more precise than traditional,

Funding: The study was funded by grants from the EORTC Quality of Life Group. The work of E.-M.G. and J.M.G. was funded by a grant from the Austrian Science Fund (FWF L502 and FWF J3353).

Conflict of interest: None.

* Corresponding author. Tel.: +45 3531 2025; fax: +45 3531 2071.
E-mail address: Morten.Aagaard.Petersen@regionh.dk (M.Aa. Petersen).

http://dx.doi.org/10.1016/j.jclinepi.2015.09.002
0895-4356/

What is new?

Key findings

- As part of the development of a computerized adaptive testing (CAT) version of the EORTC QLQ-C30 questionnaire the EORTC Quality of Life Group has developed an emotional functioning item bank of 24 items.

What this adds to what was known?

- CAT measurement based on the item bank had high levels of measurement precision and efficiency.

What is the implication and what should change now?

- The new EORTC CAT measure allows for more precise and flexible measurement of emotional functioning than the existing QLQ-C30 scale and may clearly reduce sample size requirements.

"static" questionnaires asking the same number of items and more efficient in the sense that fewer items are needed to obtain a specific precision. CAT instruments are also highly flexible as they can be adapted to the requirements of each study or setting. Because of these advantages of CAT, several groups have developed and/or explored the use of CAT to assess patient-reported outcomes (PROs) [6–13].

The European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life questionnaire (QLQ-C30) is an internationally widely used instrument for the assessment of health-related quality of life (HRQOL) in cancer patients [14,15]. It consists of 30 items measuring 15 aspects of HRQOL: five functional measures, nine symptom measures, and one measure of overall health/quality of life [16]. To improve the assessment of PROs in oncology, the EORTC Quality of Life Group is currently developing CAT versions of the EORTC QLQ-C30 scales [17–23]. The new CAT instrument operates within the same conceptual framework as the QLQ-C30. Hence, the aim is to develop a unidimensional item bank for each QLQ-C30 scale, which, in addition to the original QLQ-C30 items, consists of items covering the same aspects of the dimension as the QLQ-C30. That is, each new item bank will include all QLQ-C30 items from the relevant scale. To further enhance the compatibility with the QLQ-C30 and to ensure a homogeneous and simple format, the new items should have the same item style as the QLQ-C30 items, that is, they should use the same response options and recall period. In this way, the CAT instrument will measure some well-validated and (to many) well-known HRQOL dimensions, and it can be related to the substantial literature of studies using the QLQ-C30.

One of the key domains of the QLQ-C30 is emotional functioning (EF). The QLQ-C30 EF scale consists of four items measuring depression, anxiety (two items), and general distress that are assumed to represent a unidimensional construct. Responses to the four items are summed to form a unidimensional EF score. The new item bank should include the four QLQ-C30 EF items. Hence, the aim is a unidimensional item bank comprising the QLQ-C30 EF items and as many additional items on depression, anxiety, and general distress as possible.

As for any EORTC instrument, development of the CAT instrument takes place in an international, cross-cultural setting. The EORTC CAT development procedure consists of four phases: (I) literature search, (II) operationalization, (III) pretesting, and (IV) field testing. Phases I–III of the EF CAT development have been completed and described elsewhere [17]. In phase I, we identified 1,729 EF items from existing questionnaires. The large majority of these items (1,480) were excluded mainly because of redundancy or lack of relevance for the EORTC measurement of EF. The remaining items formed the basis for formulating new EF items fitting the "QLQ-C30 item style." After a second round of evaluations of redundancy and relevance, 63 items were retained. Evaluations by international samples of experts (phase II) and cancer patients (phase III) further reduced this to 38 candidate items. The present article reports on the phase IV field testing and psychometric evaluations of the 38 EF items.

2. Methods

The methods and analyses used in phase IV for the final development of the EF item bank are described in the following. They generally follow the approach previously reported for other dimensions [20–22]. Please refer to these publications for further details.

2.1. Sample

The EORTC CAT is intended for international use for cancer patients in general. Therefore, we accrued an international sample of cancer patients with different diagnoses, stages of disease, and so forth. Patients were recruited from oncology departments in Austria, Denmark, Italy, and the United Kingdom in the period February to December 2011. Patients were invited either by mail or when coming to the department. Eligibility requirements included a verified cancer diagnosis, age at least 18 years, and being physically and mentally competent to complete the questionnaire. Written informed consent was obtained following local requirements.

The study was approved by the local ethics committees of the participating countries.