Interpretability of change in the Nurses Work Functioning Questionnaire: minimal important change and smallest detectable change

Fania R. Gärtner, Karen Nieuwenhuijsen*, Frank J.H. van Dijk, Judith K. Sluiter

Coronel Institute of Occupational Health, Academic Medical Center, University of Amsterdam, PO Box 22700, 1100 DE Amsterdam, The Netherlands

Abstract

Objectives: The present study evaluates the interpretability of individual changes and assesses the minimal important change (MIC) for improvement and the smallest detectable change (SDC) of the Nurses Work Functioning Questionnaire (NWFQ).

Study Design and Setting: Data of two time points, baseline and 3-month follow-up, of a randomized controlled trial were used. A total of 358 nurses and allied health professionals of one Dutch academic medical center participated at both measurement points. Separate MIC values were calculated for the total score of the NWFQ and its six subscales, using two anchor-based methods: mean change and receiver operating characteristics (ROC) curve methods. Two methods for baseline corrections were applied: subgroup analyses and MIC calculation based on relative change scores. The SDCs were calculated using the standard error of the measurement.

Results: MIC values ranged from 3.4 to 8.3 for the mean change method and from 1.5 to 9.5 for the ROC curve method. In a subgroup with high-baseline scores, the MIC values of the two methods ranged from 4.4 to 29 and 9.5 to 41.5, respectively. The SDC values ranged from 7.2 to 17. Only one MIC value exceeded the SDC; however, 10 of the 14 MIC values exceeded the SDC in the high-baseline group.

Conclusion: Three of the seven NWFQ scales exhibited sufficient interpretability of individual change. For four scales, conclusions on the interpretability of change cannot yet be drawn. SDCs were small compared with the scale range. © 2012 Elsevier Inc. All rights reserved.

Keywords: NWFQ; Allied health professionals; Common mental disorders; Occupational health; Presenteeism; Clinimetric quality

1. Introduction

Mental health illness impairs the work functioning of workers who maintain working during complaints [1–5]. These impairments in work functioning because of mental health complaints may have serious consequences in some occupations, such as the health care service, with risks for patient safety. Effective interventions that address the mental health complaints and its effect on work functioning would benefit employers and workers. For this aim, a few interventions have been developed and evaluated over the last decade. One example is an intervention study of an adjuvant occupational therapy in workers with depression [6].

The efficacy of interventions in the enhancement of work functioning has primarily been evaluated using self-reported outcome measures. However, systematic reviews reveal that many existing instruments of work functioning have shortcomings in quality evaluation, especially in the interpretability of individual change, defined as the ability to detect change in the measured construct over time

[7,8]. The Nurses Work Functioning Questionnaire (NWFQ) is a newly developed measurement scale for the assessment of work function impairments because of mental health complaints [9]. The NWFQ is a 47-item self-report questionnaire with six subscales that measure various aspects of work function impairments and was specifically developed for nurses and allied health professionals. A recent study revealed that the NWFQ is a valid and reliable measurement scale [10]. However, for application in intervention studies and practice, the interpretability of change requires further investigation to facilitate the interpretation of individual change scores [11]. The statistical significance of differences alone does not provide insight on whether changes in outcome scores are "real" changes or changes that occur because of measurement error. Moreover, the clinical relevance of real changes remains questionable. The interpretability of change includes two measures, the "minimal important change" (MIC) and the "smallest detectable change" (SDC). The MIC value refers to the smallest difference between two scores on a measurement scale that can be regarded as relevant or important [12,13]. The SDC value refers to the smallest change that can be regarded as a real change that is not because of measurement

* Corresponding author. Tel.: +31-20-5665320; fax: +31-20-6977161.
E-mail address: K.Nieuwenhuijsen@amc.uva.nl (K. Nieuwenhuijsen).

What is new?

- The present study describes the evaluation of the interpretability of individual change and the calculation of the minimal important change (MIC) and smallest detectable change (SDC) for the total score of the Nurses Work Functioning Questionnaire (NWFQ), a newly developed measurement scale for the measurement of work functioning impairments, and its subscales. The MIC values of the NWFQ total scale and its subscales have been calculated as well as its SDC scores.

- The present study presents an illustration of the methodology for the evaluation of interpretability and the calculation of the MIC—methods that have recently undergone a fast development.

- The knowledge of SDC and MIC is crucial for the interpretation of individual change scores of the NWFQ. These values might guide researchers and practitioners in their conclusions on whether changes in individual workers are ''real'' and ''relevant.''

- The formulation of an accurate anchor question, a one-item global rating scale asking for perceived change in work functioning, which was used as a reference standard for the calculation of MIC, appears complex. Further research is needed to identify a more suitable formulation for anchor questions on perceived change in work functioning.

error [11,14]. With these two measures, we have an indication on whether the SDC is small enough to detect MICs.

Both values aid in the interpretation of results of individual subjects for the scientific evaluation of interventions. Furthermore, occupational health practitioners who advise nurses with mental health complaints would receive indications on relevant changes in the work functioning of their patients. Therefore, this study will assess the MIC and SDC of NWFQ total scores and each of the different subscales.

The methodology for MIC calculation has been in development since its introduction in the late 1980s. Jaeschke et al. [15] first introduced a definition and method for MIC calculation. Various review articles on potential methods have been published previously [11–13,16–21], but no consensus on the best methodology has been reached. MIC calculation is possible for changes in improvement and deterioration. This study will focus on the MIC for improvement, which from this point will be referred to as MIC. Two main approaches have been described in the literature: the distribution-based methods [22–24] and the anchor-based methods [15,25]. As Terwee et al. [21] defines,

distribution-based methods use statistical measures as a value for MIC and anchor-based methods use an external criterion, often called ''anchor,'' to define an important change. However, based on a comparison of these two methods, Terwee et al. [21] conclude that from a conceptual point of view, the anchor-based methods should be preferred above the distribution-based methods. We agree with the preference for anchor-based methods because in our view, anchors represent perceptions of subjects on relevant changes, whereas purely statistical methods are barely possible to provide answers on relevance. Therefore, we will apply the anchor-based methods in this study, in which an external standard is used for the MIC calculation using a self-report measure on the perception of an important change as an anchor, as is recommended by Crosby et al. [17]. As several MIC calculation methods are available in the anchor-based approach, we applied the triangulation of two methods, the mean change and the receiver operating characteristics (ROC) curve methods.

Another issue is the baseline value of subjects. Two methods are mentioned in the literature to take into account differences in the baseline values: the analysis of subgroup MIC values and the calculation of the MIC based on relative change scores [17,18,21,26]. Both methods will be applied in the present study because no consensus on a preferred method has been reached.

The first objective of this study is the evaluation of good interpretability of individual change scores by calculating the area under the curve (AUC) of the ROC curve and by comparing the resulting MIC values with the SDC value. The second objective is to calculate MIC and SDC for the interpretation of individual change scores in research and practice. These values will be calculated for the NWFQ total score and for the separate NWFQ subscales.

2. Methods

2.1. Design

This study used a repeated measures design with a baseline (T_0) and a 3-month follow-up measure (T_1). Data collection was integrated in the Mental Vitality @ Work study, a cluster randomized controlled trial (RCT) with three arms: one control group and two intervention groups. The RCT examines the effect of a workers' health surveillance mental module for nurses and allied health professionals. Participants in both intervention arms completed an online screening questionnaire to detect problems in mental health and work functioning and received immediate feedback on their screening results. A consultation with an occupational physician was offered in cases of impairments in mental health or work functioning in the first intervention arm. The second intervention arm offered self-help e-mental health interventions. The study design and procedure have been described previously by Gärtner et al. [27]. The data