ELSEVIER

Journal of Clinical Epidemiology 64 (2011) 301−308

Journal of Clinical Epidemiology

Validation testing of a three-component model of Short Form-36 scores

Yoshimi Suzukamo[a,b,*], Shunichi Fukuhara[a,c], Joseph Green[a,d], Mark Kosinski[e], Barbara Gandek[e], John E. Ware[f]

[a]*Institute for Health Outcomes and Process Evaluation Research, Kyoto and Tokyo, Japan*
[b]*Department of Physical Medicine and Rehabilitation, Tohoku University, Sendai, Japan*
[c]*Department of Epidemiology and Healthcare Research, Kyoto University, Kyoto, Japan*
[d]*Graduate School of Medicine, The University of Tokyo, Tokyo, Japan*
[e]*QualityMetric Inc., Lincoln, RI, USA*
[f]*University of Massachusetts-Boston, Boston, MA, USA*

Accepted 13 April 2010

Abstract

Objective: The two-component factor structure underlying Short Form-36 (SF-36) summary scores may not be valid worldwide. We studied a three-component model of SF-36 scores in Japan.

Study Design and Setting: The SF-36 scores came from representative samples of the population of Japan. Factor analysis and structural equation modeling were used. The two-component model gave physical component summary (PCS) scores and mental component summary (MCS) scores. The three-component model gave scores on the PCS, the MCS, and also on the third component, which we call the role component summary (RCS) score. These were evaluated with external criteria.

Results: In the three-component model, the RCS was strongly associated with the role-physical, social functioning, and role-emotional subscales, whereas the PCS and MCS were associated with the physical functioning and mental health subscales, as expected. The goodness-of-fit index was 0.945 for the three-component model and 0.935 for the two-component model. The PCS discriminated between groups stratified by comorbid conditions, and the MCS discriminated between groups stratified by psychological depression. Absence from work was associated with both PCS and RCS.

Conclusion: The three-component model is better than the two-component model, and it provides more useful PCS and MCS scores. Criteria for validation testing of the RCS are needed.

Keywords: Quality of life; Structure; SF-36; Summary score; Role/social; Validity

1. Introduction

Measures of health-related quality of life are often based on explicit conceptual models. The model associated with the original, US-English version of the Short Form-36 (SF-36) has eight subscales [1,2]. The four subscales measuring physical functioning (PF), limitations on role functioning because of physical health (RP), bodily pain (BP), and general health (GH) are often given the greatest weights to form a ''physical component.'' The other four, which measure mental health (MH), limitations on role functioning because of emotional problems (RE), social functioning (SF), and vitality (VT) are the greatest contributors to a ''mental component'' [3]. ''The physical and mental components are conceived as being parts of a higher-order concept: health-related quality of life'' [1,4]. Through the International Quality of Life Assessment project [5], translated and adapted versions of the SF-36 were made available for use in many countries and language-based cohorts. The SF-36 has been used worldwide, and results from studies done in East Asia have now brought its original conceptual model into question.

Factor structures differing from that found in Western countries have been found in Japan [6], China [7,8], Singapore [9], and Taiwan [10], although not in Hong Kong [11]. In Japan, Fukuhara et al. [6] found that the factor structure of the SF-36 differed from that in other countries in three ways: (1) Scores on the role-emotional (RE) subscale loaded strongly on the ''mental'' component in the United States and in Western European countries, but they loaded strongly on the ''physical'' component in Japan. (2) Scores on the VT subscale loaded strongly on the ''physical'' component in the other countries, but they loaded strongly on the

* Corresponding author. Department of Physical Medicine and Rehabilitation, Tohoku University Graduate School of Medicine, 2-1 Seiryo-cho, Aoba-ku, Sendai 980-8575, Japan. Tel.: +81-22-717-7338; fax: +81-22-717-7340.

E-mail address: suzukamo@med.tohoku.ac.jp (Y. Suzukamo).

0895-4356/$ - see front matter
doi: 10.1016/j.jclinepi.2010.04.017

What is new?

For SF-36 scores in Japan, using a model that includes three components (role-social, physical, and mental) solves problems that arise when a simpler model is used.

''mental'' component in Japan. (3) Scores on the BP subscale loaded strongly on the ''physical'' component in the other countries, but they loaded strongly on both the ''mental'' and the ''physical'' components in Japan.

The original SF-36 measurement model has also been reconsidered in the light of findings from Western countries. Using structural equation modeling and data from the United States and Europe, Keller et al. [4] found support for explanations invoking a ''general well-being'' factor at the same level as the ''physical'' and ''mental'' factors, and also a single higher-level factor (interpreted as ''health''). Five years later, Ware [12] proposed a model including a ''participation (role, social)'' factor in addition to the ''physical'' and ''mental'' factors.

In addition to the eight subscale scores, physical component summary (PCS) scores and mental component summary (MCS) scores are often computed, at least in Western countries. As convenient as these summary scores may be, their validity depends on the appropriateness of the two-factor model. The differences in factor structure between Western and Asian countries, and appreciation of the value of role and social participation as an independent domain, have led to proposals for three-component summaries of SF-36 scores. For example, using structural equation modeling, Huang et al. [13] found a good fit between SF-36 data from the general population of Taiwan and a model with three second-order components (which they interpreted as physical, mental, and social) and one third-order component (which they interpreted as health).

To develop a more appropriate way to summarize SF-36 scores, we compared a model with two components (mental and physical) to a model with three components (mental, physical, and role-social).

2. Methods

2.1. Sample

To study the factor structure of the SF-36 scores and to determine how to compute the three summary scores, we used data from the 2002 national norm survey done in Japan (sample 1, $n = 2{,}966$) [14]. For validation tests of the three-component model, we used data from the 1995 national norm survey [6] (sample 2, $n = 3{,}395$; some data used for validation tests were missing in 308 cases; hence, the total number used was 3,087). All residents of Japan who were 20 through 80 years old were potential subjects of these surveys, and participants were selected by two-stage stratified random sampling. All participants completed the SF-36 self-report form.

2.2. Data collection

The 2002 national norm survey was done with the Japanese SF-36 version 2, and the 1995 survey was done with version 1.2. For six of eight subscales, the difference in scores between versions was less than 1.0 point [14]. For the scales measuring limitations on role functioning because of physical health and emotional problems, the differences were somewhat greater, as in the United States [15]. The SF-36 comprises one question item about recent changes in health (the data from which were not used in this study) and 35 question items that are scored in eight subscales: PF, RP, BP, GH, VT, SF, RE, and MH.

Together with the SF-36, the respondents were given a list of 19 medical conditions and, for each condition, were asked to indicate whether they had the condition or not. The 19 medical conditions were hypertension, diabetes, stroke, myocardial infarction, angina, congestive heart failure, eye disease, respiratory disease, gastrointestinal disease, blood disease, kidney disease, urological disease, bone or muscle disease, skin disease, neurological disease, depression, hormonal disease, gynecological disease, and any other chronic condition.

Respondents were also asked to complete the Zung Self-Rating Depression Scale (ZSDS) [16]. Reliability and validity studies of the Japanese version of the ZSDS have been reported [17], and scores on the ZSDS (lowest possible score, 20; highest possible score, 80) were taken as indicators of the severity of depression, as follows: greater than 55, severe; 48 through 55, moderate; and 40 through 47, mild.

In addition, the respondents were asked to indicate the number of days they had missed work, school, or housework for health-related reasons during the previous year.

2.3. Analysis

The original measurement model is a two-component subscale-level model [1−3]. To test the hypothesis of a three-component structure, we first used exploratory factor analysis. With the limit set at three, factors were extracted using the principal components method, followed by varimax rotation, which resulted in the three-component subscale-level model. The results were examined to see whether they fulfilled the following criteria:

1. At least 70% of the variance should be explained by the three factors.
2. The subscale that correlated most strongly with the physical component should be the PF subscale.
3. The subscale that correlated most strongly with the mental component should be the MH subscale.
4. The subscales that correlated most strongly with the role-social component should be the RP, RE, and SF subscales.