ELSEVIER

Journal of Clinical Epidemiology 58 (2005) 815–822

Journal of Clinical Epidemiology

The SF-36 summary scales were valid, reliable, and equivalent in a Chinese population

Cindy L.K. Lam[a,*], Eileen Y.Y. Tse[a], Barbara Gandek[b], Daniel Y.T. Fong[c]

[a]*Family Medicine Unit, University of Hong Kong, 3rd Floor, Ap Lei Chau Clinic, 161 Main Street, Ap Lei Chau, Hong Kong SAR, China*
[b]*IQOLA Project, Health Assessment Lab, Waltham, MA, USA*
[c]*Department of Nursing Studies, University of Hong Kong, Hong Kong SAR, China*

Accepted 8 December 2004

Abstract

Objectives: To find out whether the SF-36 physical and mental health summary (PCS and MCS) scales are valid and equivalent in the Chinese population in Hong Kong (HK).

Study Design and Setting: The SF-36 data of a cross-sectional study on 2,410 Chinese adults randomly selected from the general population in HK were analyzed.

Results: The hypothesized two-factor structure of the physical and mental health summary scales (PCS and MCS) was replicated and the expected differences in scores between known morbidity groups were shown. The internal reliability coefficients of the PCS and MCS scales ranged from 0.85 to 0.87. The effect size differences between the U.S. standard and HK-specific PCS and MCS scores were mostly <0.5. The effect size differences in the standard PCS and MCS scores of specific groups between the U.S. and H.K. populations were all <0.5.

Conclusion: The PCS and MCS scales were applicable to the Chinese population in HK. The high level of measurement equivalence of the scales between the U.S. and H.K. populations suggests that data pooling between the two populations could be possible. To our knowledge, this is the first study to show that the SF-36 summary scales are valid and equivalent in an Asian population.

Keywords: Health-related quality of life; SF-36 summary scales; Construct validity; Reliability; Cross-cultural equivalence; Chinese

1. Introduction

Health-related quality of life (HRQoL), defined by Bullinger et al. [1] as "the impact of perceived health on an individual's ability to live a fulfilling life," is becoming an important outcome measure in health services and clinical trials. The MOS 36-item Short-Form Health Survey (SF-36) is a popular HRQoL measure that has been translated and validated for Chinese adults in Hong Kong (HK) [2–5]. The SF-36 has eight scales measuring eight domains of HRQoL: physical functioning (PF); role–physical (RP), or limitation in daily role functioning due to physical problems; role–emotional (RE), or limitation in daily role functioning due to emotional problems; bodily pain (BP); general health perception (GH); vitality (VT); social functioning (SF); and mental health perception (MH). Each scale consists of 2 to 10 items, and each item is rated on a two- to six-point Likert scale. The scale score is calculated by summation of all the scores of items belonging to the same scale. A profile of eight scale scores, although informative, can be difficult to interpret as an outcome measure in clinical trials [6]. Ware et al. [6–8] hypothesized that there are two principal factors, namely the physical and the mental components, underlying the eight SF-36 scales. This two-factor structure was demonstrated in the general population in the United States (U.S. standard): the physical health summary (PCS) and mental health summary (MCS) components explained 60% of the total variance of the SF-36 scale scores [6–8]. The physical component correlated strongly ($r \geqslant .7$) with the physical functioning (PF), role–physical (RP), and bodily pain (BP) scales but weakly ($r \leqslant .3$) with the mental health (MH), role–emotional (RE), and social functioning (SF) scales. The mental component correlated strongly with the MH, RE, and SF scales but weakly with the PF, RP, and BP scales. The general health (GH) and vitality (VT) were bipolar scales, loading moderately ($.3 < r < .7$) on both physical and mental components [6–8].

The PCS and MCS scales summarize the eight SF-36 scale scores into two summary scores that give an overall assessment of quality of life related to physical and

* Corresponding author. Fax: +852-28147475.
E-mail address: clklam@hku.hk (C.L.K. Lam).

0895-4356/05/$ – see front matter
doi: 10.1016/j.jclinepi.2004.12.008

mental health, respectively. The PCS and MCS scores are easier to interpret and simpler to analyze statistically in clinical trials and longitudinal studies [6,7]. Because different SF-36 scales correlate with each of the two factors differently, they are weighted by the appropriate physical or mental factor coefficients before aggregation to form the two summary scores. Norm-based scoring with *z*-score transformation, calculated as (observed score – population mean)/population standard deviation, and standardization of the population mean and standard deviation (SD) to 50 and 10, respectively, are recommended for easier interpretation [6]. The SF-36 PCS and MCS scoring algorithm is summarized below:

$$\text{SF-36 PCS} = \sum(z\text{-score of each scale} \times \text{respective physical factor coefficient}) \times 10 + 50$$

$$\text{SF-36 MCS} = \sum(z\text{-score of each scale} \times \text{respective mental factor coefficient}) \times 10 + 50$$

The standard SF-36 PCS and MCS scales scoring algorithm uses the population means, SD, and factor coefficients derived from the U.S. general population [6]. A multinational study showed similar factor structures and equivalent population mean PCS and MCS scores between the United States and nine European countries [8,9]. Ware et al. [8] recommended that the U.S. standard SF-36 PCS and MCS scales and scoring algorithm should be used in these countries, instead of country-specific approaches. Data from the Japanese general population, however, and from several Chinese populations, showed the two principal factor structure and loadings of the SF-36 scales differing from those found in the U.S. population [10–13]. These studies found that the role–emotional scale loaded more strongly (r = .62–.82) on the physical than the mental component (r = .19 to .49), which was the reverse of that found in the U.S. data (physical: r = .17, mental: r = .78). The vitality scale loaded strongly (r = .79–.88) on the mental component but only weakly (r = .21–.37) on the physical component in these populations, instead of the moderate correlations with both components found in the U.S. data (physical: r = .47, mental: r = .64). This raised a concern of whether the standard PCS and MCS scales are applicable to Asian populations, whose cultures may differ more than the European cultures from that of the United States.

Our objective was to find out whether the SF-36 PCS and MCS scales are valid, reliable, and equivalent for the H.K. Chinese adult population. We also wanted to find out whether a HK-specific scoring algorithm using factor coefficients derived from the H.K. general population would give results equivalent to those of the standard algorithm. Evidence on validity and reliability would support the use of the SF-36 PCS and MCS scales in HK. Equivalence in results between the U.S. and H.K. Chinese populations implies that the standard SF-36 PCS and MCS scales can be used as a cross-cultural HRQoL measure in international studies and global drug trials [14].

2. Methods

Data of 2,410 Chinese adults randomly selected from the general population in HK that were collected in a cross-sectional norming study of the Chinese (Hong Kong) SF-36 Health Survey in 1998 were used. The detailed sampling and data collection methods have been described elsewhere [3,5]. The sociodemographic characteristics of the subjects are compared to those of the H.K. general adult population in Table 1.

The data were tested against the following hypotheses.

1. Two principal component factors with eigenvalues >1.0 can be extracted from the eight SF-36 scale

Table 1
Sociodemographic characteristics of H.K. Chinese study sample compared with the Hong Kong general adult population

Characteristic	H.K. sample	H.K. adults[a]
Sample size	N = 2,410	N = 5,333,610
Mean age, years	42.9	42.3
Age group, %		
18–44 years	56.7	58.6
45–64 years	23.7	27.4
⩾65 years	15.3	14.0
Refused to answer	4.2	0
Sex, %		
Male	47.8	48.3
Female	52.2	51.7
Marital status, %		
Now married	58.0	59.4
Never married	33.8	31.9
Widowed	5.8	6.0
Divorced or separated	1.3	2.7
Refused to answer	1.1	0
Education, %		
No schooling	6.9	8.4
Primary (1–6 years)	22.3	20.5
Secondary (7–13 years)	52.2	54.6
Tertiary (college and beyond)	17.8	16.4
Refused to answer	0.9	0
Social class by occupation, %		
Managers and administrators[b]	n/a	10.7
Professional	3.1	5.5
Associate professional	14.7	15.0
Skilled workers[c]	35.4	33.5
Semiskilled workers[d]	24.6	15.0
Nonskilled workers[e]	14.4	19.8
Refused to answer	7.7	0

Abbreviations: n/a, not available.

[a] Data from the Hong Kong 2001 population census.

[b] This occupation category is not applicable to the social class by occupation classification.

[c] Craft workers, plant and machine operators, and assemblers.

[d] Service and shop sales workers.

[e] Workers in primary industries, including agriculture and fisheries, and unclassified.