The SF-36 Summary Scales is Valid, Reliable and Equivalent in a Chinese Population

Cindy L.K. Lam, MBBS, MD, FRCGP, FHKAM (Family Medicine)

Associate Professor, Family Medicine Unit, the University of Hong Kong

Eileen Y. Y. Tse, MBChB, FRCGP, FHKAM (Family Medicine)

Assistant Professor, Family Medicine Unit, the University of Hong Kong

Barbara Gandek, MS

Scientist, Health Assessment Lab, Waltham, USA

Daniel Y T Fong, Ph D,

Assistant Professor, Department of Nursing Studies, The University of Hong Kong

Correspondence Address: Family Medicine Unit, The University of Hong Kong, 3rd Floor, Ap Lei Chau Clinic, 161 Main Street, Ap Lei Chau, Hong Kong SAR. Fax: 852 28147475; E mail: clklam@hku.hk

The SF-36 Summary Scales is Valid, Reliable and Equivalent in a Chinese Population

Abstract

Objectives: To find out whether the SF-36 Physical and Mental Health Summary (PCS and MCS) Scales were valid and equivalent in the Chinese population in Hong Kong (HK) or not. **Methods**: The SF-36 data of a cross-sectional study on 2410 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 (PCS and MCS) Scales 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 standard (US) and HK specific PCS and MCS scores were mostly less than 0.5. The effect size differences in the standard PCS and MCS scores of specific groups between the US and HK populations were all less than 0.5. **Conclusion**: The PCS and MCS Scales were applicable to the Chinese population in Hong Kong. The high level of measurement equivalence of the Scales between the US and HK populations suggests that data pooling between the two populations could be possible. This was the first study to show that the SF-36 Summary Scales were valid and equivalent in an Asian population.

Key words: Health-related quality of life, SF-36 Summary Scales, Construct validity, reliability, Cross-cultural Equivalence, Chinese

Running Title: SF-36 PCS & MCS in a Chinese Population

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 health-related quality of life measure (HRQOL) that has been translated and validated for Chinese adults in Hong Kong [2-5]. The SF-36 has eight scales measuring eight domains of HRQOL: physical functioning (PF); role-physical (RP), i.e. limitation in daily role functioning due to physical problems; role-emotional (RE), i.e. 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 two to ten 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 hypothesized that there were two principal factors, namely the physical and the mental components, underlying the eight SF-36 scales [6-8]. This two-factor structure was demonstrated in the general population in the United States (US); 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 ($r \geq 0.7$) with the physical functioning (PF), role-physical (RP) and bodily pain (BP) scales but weakly ($r \leq 0.3$) with the mental health (MH), role-emotional (RE) and social functioning (SF) scales, while 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 that loaded moderately ($0.3 < r < 0.7$) on both physical and mental components [6-8].

The PCS and MCS Scales summarise the eight SF-36 scale scores into two summary scores that give an overall assessment of quality of life related to physical and mental health, respectively. The PCS and MCS scores are easier to interpret and simpler to analyse statistically in clinical trials and longitudinal studies [6,7]. Since 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 [(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: -

SF-36 PCS = Σ(z score of each scale x respective physical factor coefficient) x 10 + 50

SF-36 MCS = Σ(z score of each scale x respective mental factor coefficient) x 10 + 50

The standard SF-36 PCS and MCS Scales scoring algorithm uses the population means, standard deviations, and factor coefficients derived from the US general population [6]. A multinational study showed similar factor structures and equivalent population mean PCS and MCS scores between the US and nine European countries [8,9]. Ware et al recommended that the standard (US), instead of country-specific, SF-36 PCS and MCS Scales and scoring algorithm should be used in these countries. However, Data from the Japanese general population and several Chinese populations showed that the two principal factor structure and loadings of the SF-36 scales were different from those found in the US population [10-13]. These studies found that the role-emotional scale loaded more strongly (r = 0.62 to 0.82) on the physical than the mental component (r = 0.19 to 0.49), which was reverse of that found in the US (physical: r = 0.17, mental: r = 0.78). The vitality scale

loaded strongly (r = 0.79 to 0.88) on the mental component but only weakly (r = 0.21 to 0.37) on the physical component in these populations instead of the moderate correlations with both components found in the US (physical: r =0.47, mental: r =0.64). This raised a concern of whether the standard PCS and MCS Scales were applicable to Asian populations whose cultures may differ more than the European cultures from that of the US.

The aim of this study was to find out whether the SF-36 PCS and MCS Scales were valid, reliable and equivalent for the Chinese adult population in Hong Kong (HK) or not. We would also like to find out whether a HK specific scoring algorithm using factor coefficients derived from the HK general population would give equivalent results as those of the standard algorithm. Evidence on validity and reliability would support the use of the SF-36 PCS and MCS Scales in Hong Kong. Equivalence in the results between the US and HK 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].

Methods

Data of 2410 Chinese adults randomly selected from the general population in Hong Kong 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 in previous papers [3,5]. The sociodemographic characteristics of the subjects are compared to those of the HK general adult population in Table 1.

The data were tested against the following hypotheses: -

1. Two principal component factors with eigenvalues greater than 1.0 can be extracted from the eight SF-36 scale scores, which can explain at least 60% of the total variance of the SF-36 scores, and ≥ 70% of the reliable variance of each scale score [6,8].
2. The two-factor structure and loadings extracted from the SF-36 scale scores of the Chinese adult population in Hong Kong are similar to those of the US population.
3. The internal reliabilities of the PCS and MCS Scales are comparable to those of the individual SF-36 Scales, and the coefficients would be greater than 0.7 [6].
4. The PCS and MCS scores of people with chronic diseases are lower than those of people without any chronic disease, the PCS score should be reduced more than the MCS score in patients with a predominantly physical condition such as heart disease, and the MCS score should be reduced more than the PCS score in patients with a psychological condition [6,7].
5. The correlation between the PCS and MCS scores should be very weak (close to zero) because they measure separate concepts [8,9].
6. The PCS and MCS scores calculated by the HK specific and the standard algorithms should have very strong correlations (>0.90) and measurement equivalence [9].

7. The standard SF-36 PCS and MCS Scales have measurement equivalence between the US and HK Chinese populations.

The standard SF-36 PCS and MCS scores were calculated by the scoring algorithm described in the SF-36 Physical and Mental Health Summary Scales User's Manual [6]. The HK specific SF-36 PCS and MCS scores were calculated with the means, SD and factor coefficients derived from the HK general Chinese population sample. Correlations between PCS and MCS scores were determined by Pearson Correlation Tests.

The internal reliability of the standard and HK specific PCS and MCS Scales were calculated by the method recommended by Ware et al, which takes into account the internal reliability of each scale and the covariances among them [6].

The standard and HK specific mean SF-36 PCS and MCS scores were calculated for the whole sample, and by chronic disease groups. The presence of chronic disease was defined by self-reporting. Each subject was asked if he/she had ever been diagnosed by a doctor to have hypertension, diabetes mellitus, heart disease, stroke, chronic pulmonary disease, chronic joint problem, psychological illness, or any other chronic disease. The significance of the differences between each disease group and the no-disease group was tested by two-sample t test and a p value ≤ 0.05 was considered statistically significant.

Scores were considered to have measurement equivalence if the effect size difference was less than 0.5, which has been recommended by many to be the minimally important difference (MID) [15-17]. The effect size difference between the standard and HK specific PCS and MCS scores was calculated by dividing their difference by the SD of the HK specific summary score. The mean standard SF-36 PCS and MCS scores of subjects with hypertension, diabetes mellitus, heart diseases and psychological problems were compared to the mean scores of corresponding disease groups from the US general population [6]. The

effect size difference in the standard PCS and MCS scores between US and HK patients was calculated by dividing their differences by the SD of the HK disease group.

All data analyses were carried out by the SPSS Programme for Windows 10.0.

Results

The Hong Kong Specific SF-36 PCS and MCS Scales

Two principal component factors were extracted from the eight SF-36 scale scores and the eigenvalues were 3.4968 and 1.1118 for the first two components, respectively. The two principal factor structure and factor loadings, after varimax rotation, of the SF-36 scale scores of the HK Chinese adult population is shown in Table 2. The physical (first) component correlated more strongly with the physical functioning (PF), role-physical (RP) bodily pain (BP) and general health (GH) than other scales, while the mental component correlated more strongly with the mental health (MH), role-emotional (RE) and social functioning (SF) than other scales. The vitality (VT) scale loaded moderately on both physical and mental (second) components. The results were largely similar to those found in the US population, which are shown in brackets in Table 2.

The factor coefficients and the variances of each scale explained by the two factors in the Hong Kong Chinese population are shown in Table 3. The standard factor coefficients derived from the US general population are also shown in brackets for comparison. The two factors explained 57.6% of the total variance of the scale scores. They explained 63% to 88% of the reliable variance of each scale. The scoring algorithm of the HK specific SF-36 MCS and PCS Scales derived from the SF-36 scale means, standard deviations and factor coefficients of the Hong Kong general population is shown in Table 4.

Construct Validity and Reliability of the SF-36 PCS and MCS Scales

The correlations between the corresponding standard and HK specific SF-36 summary scores were close to unity (0.975 for the PCS and 0.985 for the MCS). The correlations between the PCS and MCS scores were zero for the HK specific scores and –0.126 for the standard scores.

The means, standard deviations of the standard and HK specific SF-36 PCS and MCS of the whole sample and different self-reported morbidity groups are shown in Table 5. The summary scores of people with any type of chronic disease were all significantly lower than those of people without any chronic disease. Heart disease, hypertension and diabetes mellitus were associated with significantly lower PCS scores but an insignificant reduction in the MCS score. The PCS and MCS of people with psychological diseases were both significantly lowered, compared with people without any chronic disease.

The internal reliability coefficients of the standard PCS and MCS Scales were 0.85 and 0.86, respectively. The internal reliability coefficients of the HK specific PCS and MCS Scales were 0.85 and 0.87, respectively.

Equivalence of the Standard PCS and MCS Scales in Hong Kong Chinese

The mean standard PCS and MCS for the overall Hong Kong Chinese adult population were 52.83 and 47.18, respectively. They were very close to the US general population means of 50. The difference between corresponding standard and HK specific mean PCS and MCS scores ranged from 1.7 to 7.66 points for different morbidity groups. The effect size differences were all less 0.5, except for the PCS of subjects reporting to have heart diseases (Table 5).

Table 6 compares the standard SF-36 PCS and MCS scores of the HK Chinese population with those of the US general population by four chronic disease groups. The differences in scores ranged from 0.8 to 5.3, and the effect size differences were all less than 0.5.

Discussion

Construct Validity and Reliability of the SF-36 Physical and Mental Health Summary Scales

The hypothesized two principal factor structure of the SF-36 scales was replicated in the general Chinese population in Hong Kong, and the factor loadings were similar to those found in the US population [6,7]. The physical factor loading in the general health (GH) Scale was relatively stronger than hypothesized but similar to that found in the US population. This confirmed the construct validity of the internal factor structure of the SF-36 PCS and MCS Scales for the Hong Kong Chinese population. The eight Chinese (Hong Kong) SF-36 scales scores can be summarised into two (physical and mental health) summary scores.

This was the first study that replicated the original two-factor structure of the standard PCS and MCS Scales on a general population in Asia. The SF-36 scale factor structure and loadings from three Asian general populations are compared to those of the Chinese population in Hong Kong found in this study in Table 7. Major differences in the factor structure and loadings are found in the RE and VT scales between the Hong Kong population and other Asian populations. Thumboo et al found that the role-emotional (RE) scale loaded more strongly on the physical component than the mental component and the vitality (VT) scale loaded strongly on the mental factor but weakly on the physical factor in the general Chinese population in Singapore [12]. Similar deviations in the factor structure were also found in the general population in Japan [11]. The RE scale loaded moderately on both the physical and mental components, and the VT scale loaded strongly on the mental component but weakly on the physical component in the Chinese population in Taiwan [18]. The factor structure and loading of the Chinese population in Hong Kong was more similar to those of

Caucasian populations in the US and Europe than those of Asian populations [9]. A possible explanation is the westernization of the life and culture in Hong Kong as a result of British governance for more than 150 years from 1842 to 1997. Westernization might also be the reason why bilingual Chinese in Singapore and the US had factor structure and loadings similar to those of the US general population instead of other Chinese people in the same countries [19, 20]. It is interesting that Chinese people in different places [13, 19, 20], and even within the same society [12,19] can have different factor structures for the same HRQOL measure, therefore one cannot extrapolate the results from one population to another without empirical testing.

The proportion of total variance in the scale scores explained by the two factors was very close to the expected standard of 60%, suggesting that the PCS and MCS Scales is a valid summary of the Chinese (Hong Kong) SF-36 scale scores. It must be pointed out that a significant proportion of the variance in the scores of each individual SF-36 scale was not explained, thus the summary scores may be less precise or sensitive than the best SF-36 scale if the difference or change concentrates in one domain [6].

The correlations between the corresponding standard and HK specific PCS and MCS scores were near unity, confirming convergent validity. On the other hand, the standard PCS and MCS scores were negatively correlated, and the Hong Kong specific PCS and MCS had zero correlation, indicating that the physical and mental health summary scores measure two totally different concepts, which cannot be further combined.

Construct validity of the SF-36 PCS and MCS Scales was also confirmed by known group comparison. There was a significant difference in the scores between groups with and without chronic diseases. Heart disease, hypertension and diabetes mellitus were associated with significant reductions in the PCS but not the MCS, as hypothesized for predominantly physical diseases [6,7]. Psychological diseases were associated with about

the same reduction in the PCS and MCS scores, contrary to McHorney et al's finding on subjects who were screened positive of depression [7]. Subjects in this Hong Kong study had known psychological diseases, which were likely to be more severe than those of subjects in McHorney et al's study. Other studies also found that patients with diagnosed mental disorders had impairment in all SF-36 domains [21,22]. Physical diseases or symptoms are common in patients with psychiatric diseases [23], and Chinese patients have a tendency to somatize their psychological problems [24-26]. All these factors explained why the PCS score of the HK Chinese people with psychological illnesses were lower than normal.

The internal reliability coefficients of the PCS and MCS Scales (0.85 –0.87) were generally higher than those of the individual SF-36 scales (0.65-0.83), and they were just short of the standard of 0.9 for individual assessment [27]. The reliability of the MCS Scale found in the HK population was similar to those of the US and nine European populations (0.85-0.90) but that of the PCS Scale was lower (0.85 Vs 0.90-0.94).

Equivalence of the SF-36 PCS and MCS Scales

The measurement equivalence of the standard SF-36 PCS and MCS Scales between the HK and US populations was very high. The mean scores of the general Chinese population were no more than three points different from the US population mean of 50, although the differences in the unadjusted mean SF-36 scale scores between the two populations were generally large [3]. This illustrates the advantage of norm-based scoring and summary scores over profile scale scores for cross-cultural comparisons and data pooling.

Hypertension, diabetes mellitus, heart diseases and psychological problems were used as tracer conditions for the assessment of measurement equivalence because they were common, relatively well-defined, and data from the US general population were available [6].

The SF-36 PCS and MCS scores of patients with each of these chronic diseases were largely equivalent between the HK and US populations. The difference in the PCS scores of people with heart diseases were relatively larger than those found in other groups because different case definitions for heart diseases were used in the HK and US studies [3, 6]. Scores of US subjects with self-reported angina had to be used for comparison with those of HK subjects reporting any heart diseases because US data on all heart diseases subjects were not available. It was not surprising that the US scores were lower than the HK scores because some of the HK subjects might have relatively mild heart diseases. The SF-36 PCS and MCS Scales seem to be quite robust in determining the level of HRQOL impairment associated with a specific condition across different cultures, which makes it a suitable cross-cultural measure. There is potential for pooling the SF-36 PCS and MCS Scales data from the US and HK Chinese populations in multi-centre trials, but this needs to be confirmed by further studies that include subjects from both populations.

There were little differences between the factor coefficients derived from the HK Chinese population and those of the standard algorithm derived from the US population. The greatest effect size difference between the standard and HK specific PCS and MCS scores was 0.52 found in people reporting heart diseases. The difference would be smaller if the standard PCS and MCS scores were norm-based on the HK Chinese population mean and SD.

Conclusions

The hypothesized two-factor structure of the SF-36 scales was replicated from the SF-36 data of the general Chinese population in Hong Kong, and the two factors explained 57.6% of the total variance of the SF-36 Scale scores and 63% to 88% of the reliable variance of each scale. The SF-36 PCS and MCS scores showed the expected difference between known chronic disease groups, further supporting their construct validity.

The mean standard PCS and MCS of the Hong Kong general Chinese population differed from the US population mean by no more than three points, and there was no important difference between the results obtained by the Hong Kong specific and the standard algorithms for different groups. Therefore the standard SF-36 PCS and MCS Scales scoring algorithm is recommended for the Chinese population in Hong Kong for better international comparability. The Hong Kong population mean and standard deviation can be used for the z-score transformation in the calculation of the standard SF-36 PCS and MCS scores in local studies, which will make interpretation easier by adjusting the population mean and standard deviation to 50 and 10, respectively.

The high degree of measurement equivalence in the PCS and MCS scores between the Hong Kong Chinese and US general populations suggests that data pooling between these two populations could be possible, but further studies are required to confirm this. The SF-36 PCS and MCS Scales can be used as a cross-cultural health-related quality of life measure in multi-national clinical trials.

To the best of our knowledge this was the first study showing the validity, reliability and equivalence of the SF-36 Physical and Mental Health Summary (PCS and MCS) Scales in an Asian population. Further studies are required to find out whether the SF-36 PCS and

MCS Scales can be adapted to other Asian countries. Population specificity and effect of westernization on the factor structure of HRQOL also deserves more research.

Acknowledgement

The general population norming survey of the Chinese (Hong Kong) SF-36 was funded by the Health Services Research Grant, the Government of Hong Kong SAR (HSRC #711026).

I would like to thank Alex Chan, Willis Ho, Joanna Shing, Ka-Lai Chan, Wai-Hung Yu, June Chan, Chi-Kwan Wong, Wing-Yee Lai, Yick-Lok Chan and Hing-Wai Tsang, for their help in data collection and analysis.

Parts of this paper have been submitted to the University of Hong Kong for the award of the Doctor of Medicine degree.

Conflict of Interest: Nil

References

1. Bullinger M, Anderson R, Cella D, Aaronson N. Developing and evaluating cross-cultural instruments from minimum requirements to optimal models. Quality of Life Research 1993; 2: 451-459.
2. Lam C L K, Gandek B, Ren X S, Chan M S. Tests of scaling assumptions and construct validity of the Chinese (HK) version of the SF-36 Health Survey. J Clin Epidemiol 1998; 51: 1139-1147.
3. Lam C L K, Lauder I J, Lam T P, Gandek B. Population based norming of the Chinese (HK) version of the SF-36 Health Survey. H K Pract 1999; 21: 460-470.
4. Lam C L K. Reliability and construct validity of the Chinese (Hong Kong) SF-36 for patients in primary care. H K Pract 2003; 25: 468-475.
5. Lam C L K, Fong D Y T, Lauder I J, Lam T P. The effect of health-related quality of life (HRQOL) on health service utilisation of a Chinese population. Social Science and Medicine 2002; 55 (9): 1635-1646.
6. Ware J E , Kosinski M. SF-36 Physical & Mental Health Summary Scales: A Manual for Users of Version 1. Second ed. Lincoln, Rohode Island: QualityMetric; 2001.
7. McHorney C A, Ware J E, Raczek A E. The MOS 36-Item Short Form Health Survey (SF-36), II: Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care 1993; 31: 247-63.
8. Ware J E, Kosinski M, Gandek B. The factor structure of the SF-36 Health Survey in 10 countries: results from the IQOLA Project. J Clin Epidemiol 1998; 51: 1159-1165.
9. Ware J E, Gandek B, Kosinski M, et al. The equivalence of SF-36 summary health scores estimated using standard and country-specific algorithms in 10 countries: results from the IQOLA project. J Clin Epidemiol 1998; 51: 1167-1170.

10. Fuh J L, Wang S J, Lu S R, Juang K D, Lee S J. Psychometric evaluation of a Chinese (Taiwanese) version of the SF-36 health survey amongst middle-aged women from a rural community. Quality of Life Research 2000; 9: 675-683.

11. Fukuhara S, Ware J E, Kosinski M, Wada S. Psychometric and clinical tests of validity of the Japanese SF-36 Health Survey. J Clin Epidemiol 1998; 51: 1045-1053.

12. Thumboo J, Fong K Y, Machin D, et al. A community-based study of scaling assumptions and construct validity of the English (UK) and Chinese (HK) SF-36 in Singapore. Quality of Life Research 2001; 10: 175-188.

13. Ren X S, Wang X S, Liu S, Hao X. Psychometric and clinical evaluation of a Chinese version of the SF-36 Health Survey among cancer patients in China. Quality of Life Newsletter 2003; 30: 5-6.

14. Herdman M, Fox-Rushby J, Badia X. A model of equivalence in the cultural adaptation of HRQOL instruments: The universalist approach. Quality of Life Research 1998; 7: 323-335.

15. Cohen J, The t test for measures. In Cohen J, Ed, Statistical Power Analysis for the Behavioral Sciences,. Hillsdale, New Jersey:Lawrence Erlbaum Associates; 1988: 19-74.

16. Kazis L E, Anderson J J, Meenan R F. Effect sizes for interpreting changes in health status. Med Care 1989; 27: S178-S189.

17. Norman G R, Sridhar F G, Walter S D, Guyatt G H. The relation of distribution- and anchor-based approaches in interpretation of changes in health related quality of life. Med Care 2001; 39: 1039-1047.

18. Tseng H, Lu J R, Gandek B. Cultural issues in using the SF-36 Health Survey in Asia: Results from Taiwan. Health and Quality of Life Outcomes 2003; 1:72-80.

19. Thumboo J, Fong K Y, Chan S P, et al. The equivalence of English and Chinese SF-36 versions in bilingual Singapore Chinese. Quality of life Research 2002; 11: 495-503.

20. Yu J, Coons S J, Draugalis J R, Ren X S, Hays R D. Equivalence of Chinese and US-English versions of the SF-36 Health Survey. Quality of Life Research 2003; 12: 449-457.

21. Spitzer R L, Kroenke K, Linzer M, et al. Health - related quality of life in primary care patients with mental disorders. JAMA 1995; 274: 1511-1517.

22. Hays R D, Wells K B, Sherbourne C D, Rogers W, Spitzer K. Functioning and well-being outcomes of patients with chronic general medical illnesses. Arch Gen Psychiatry 1995; 52: 11-19.

23. Phelan M, Stradins L, Morrison S. Physical health of people with severe mental illness. BMJ 2001; 322: 443-444.

24. Lam C L K. How does depression present in general practice? HKMJ 1995; 1: 225-229.

25. Kleinman A. Neurasthenia and depression: a study of somatization and culture in China. Culture Med Psych 1982; 6: 117-190.

26. Wong C K, Chan T S C. Somatic symptoms among Chinese psychiatric patients. H.K.J Mental Health 1984; 13: 5-21.

27. Nunnally J C, Bernstein R H. The assessment of reliability. In: Nunnally JC, Bernstein R H. Psychometric Theory. (3rd Ed). New York: McGraw-Hill; 1994: 248-292.

Table 1: Sociodemographic Characteristics of Study Sample Compared with the Hong Kong General Population

	Sample N=2,410	Hong Kong General Adult Population [a] N=5,333,610
Mean Age (years)	42.9	42.3
Age Group (years)		
18-44	56.7%	58.6%
45-64	23.7%	27.4%
65 or above	15.3%	14.0%
Refused to answer	4.2%	0%
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/Separated	1.3%	2.7%
Refused to Answer	1.1%	0%
Educational Level		
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	N.A.	10.7% [b]
Professional	3.1%	5.5%
Associate Professional	14.7%	15.0%
Skilled Worker	35.4%	33.5%[c]
Semi-skilled Worker	24.6%	15.0%[d]
Non-skilled Worker	14.4%	19.8%[e]
Refused to Answer	7.7%	0%

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 elementary occupation, agriculture and fishery, and unclassified.

Table 2 : Correlations (r) between the Chinese (Hong Kong) SF-36 Scales and Two Principal Components

	Hypothesised (and US) Correlations		Hong Kong Chinese Adults (N=2410)	
	Physical	**Mental**	**Physical**	**Mental**
PF	) (0.85)	∗ (0.12)	0.82	0.04
RP	) (0.81)	∗ (0.27)	0.66	0.36
BP	) (0.76)	∗ (0.28)	0.72	0.21
GH	0 (0.69)	0 (0.37)	0.71	0.27
VT	0 (0.47)	0 (0.64)	0.38	0.60
SF	0 (0.42)	) (0.67)	0.21	0.65
RE	∗ (0.17)	) (0.78)	0.08	0.78
MH	∗ (0.17)	) (0.87)	0.20	0.78

↓ *Strong* ($r \geq .70$) ; 0 *Moderate to substantial* ($.30 < r < .70$); ↓ *Weak* ($r \leq .30$)

Table 3: Principal Component Factor Coefficients of the Chinese (Hong Kong) SF-36 Scale Scores

	Hong Kong (US Standard[a]) Coefficients		Total Variance Explained	Cronbach's Alpha	Reliable[b] Variance Explained
	Physical Factor	Mental Factor			
PF	0.46095 (0.42402)	-0.22743 (-0.22999)	0.67	0.81	0.83
RP	0.27474 (0.35119)	0.01327 (-0.12329)	0.57	0.83	0.69
BP	0.35475 (0.31754)	-0.09483 (-0.09731)	0.57	0.74	0.77
GH	0.32470 (0.24954)	-0.05122 (-0.01571)	0.57	0.65	0.88
VT	0.03257 (0.02877)	0.25123 (0.23534)	0.51	0.72	0.71
SF	-0.07846 (-0.00753)	0.33064 (0.26876)	0.47	0.75	0.63
RE	-0.19399 (-0.19206)	0.44834 (0.43407)	0.61	0.82	0.74
MH	-0.12198 (-0.22069)	0.41167 (0.48581)	0.65	0.78	0.83

a. The values in brackets are the Standard SF-36 PCS and MCS factor coefficients, derived from the US general population.

b. Reliable variance = total variance explained divided by the internal reliability (Cronbach's alpha) of the scale.

Table 4: Scoring Algorithm of the Hong Kong Specific SF-36 PCS and MCS Scales

PF_Z = (PF - 91.82573) / 12.88527

RP_Z = (RP - 82.42739) / 30.97154

BP_Z = (BP - 83.97801) / 21.89251

GH_Z = (GH - 55.97759) / 20.17986

VT_Z = (VT - 60.27178) / 18.64714

SF_Z = (SF - 91.19295) / 16.56710

RE_Z = (RE - 71.65975) / 38.36354.

MH_Z = (MH - 72.78506) / 16.56739

Hong Kong Specific SF-36 PCS Score

(PF_Z x 0.46095 + RP_Z x 0.27474 + BP_Z x 0.35475 + GH_Z x 0.32470 + VT_Z x 0.03257 + SF_Z x -0.07846 + RE_Z x -0.19399 + MH_Z x -0.12198) x 10 + 50

Hong Kong Specific SF-36 MCS Score

(PF_Z x -0.22743 + RP_Z x 0.01327 + BP_Z x -0.09483 + GH_Z x -0.05122 + VT_Z x 0.25123 + SF_Z x 0.33064 + RE_Z x 0.44834 + MH_Z x 0.41167) x 10 + 50

Table 5: Standard and Hong Kong Specific SF-36 PCS and MCS Scores by Disease Groups

	Std PCS	HK PCS	Std MCS	HK MCS
All Subjects (N=2410)				
Mean	52.83	50.00	47.18	50.00
SD	7.31	10.00	9.61	10.00
Effect size[+]	0.28		0.28	
No Chronic Disease (n=1493)				
Mean	55.12	53.48	47.82	50.52
SD	5.32	6.69	8.89	9.19
Effect size[+]	0.25		0.29	
Any Chronic Disease (n=917)				
Mean	49.12**	44.34**	46.13**	49.15**
SD	8.48	11.76	10.60	11.15
Effect size[+]	0.41		0.27	
Heart Diseases (n=94)				
Mean	43.70**	36.04**	45.83[#]	49.21[#]
SD	10.41	14.75	10.42	10.93
Effect size[+]	0.52		0.31	
Psychological Diseases (n=94)				
Mean	47.78**	41.57**	39.14**	41.63**
SD	9.10	12.88	10.99	11.78
Effect size[+]	0.48		0.21	
Hypertension (n=271)				
Mean	48.06**	42.65**	47.49[#]	50.80[#]
SD	9.35	13.52	10.49	10.82
Effect size[+]	0.40		0.31	
Diabetes Mellitus (n=110)				
Mean	45.50**	39.25**	47.84[#]	51.02[#]
SD	9.95	14.19	10.69	11.18
Effect size[+]	0.44		0.28	

\+ *difference between standard and Hong Kong specific mean summary score/ SD of the HK specific summary score*

** *Difference between no chronic disease and disease group is significant by the two sample t tests with p<0.01*

\# *Difference between no chronic disease and disease group is not significant by the two sample t tests with p>0.05*

Table 6: Standard SF-36 PCS and MCS Scores of Hong Kong Chinese and US Populations by Disease Groups

	Mean (SD) PCS	Mean (SD) MCS
Heart Diseases		
HK Chinese (n=94)	41.66 (10.97)	46.86 (10.21)
US[a] (n=107)	36.36 (12.38)	48.04 (12.42)
Effect size[+]	0.48	0.12
Psychological Diseases		
HK Chinese (n=110)	45.79 (9.57)	40.23 (10.77)
US[b] (n=881)	47.92 (11.62)	43.46 (11.42)
Effect size[+]	0.22	0.30
Hypertension		
HK Chinese (n=271)	46.32 (9.89)	48.41 (10.26)
US (n=670)	44.57 (11.29)	49.24 (10.55)
Effect size[+]	0.18	0.08
Diabetes Mellitus		
HK Chinese (n=110)	43.63(10.53)	48.79 (10.44)
US (n=145)	39.30 (11.32)	47.90 (11.37)
Effect size[+]	0.41	0.09

+ difference in mean summary score between H K and US disease groups / SD of the HK disease group

a. US general population subjects with self-reported diagnosis of angina.

b. US general population subjects screened positive of depressive symptoms.

Table 7: Correlations (r)* between the SF-36 Scales and Two Principal Components in Asian Populations

	Japanese [11] (N=3395)		Singapore Chinese [12] (N=1381)		Taiwan Chinese [18] (N=1191)		Hong Kong Chinese (N=2410)	
	Physical	**Mental**	**Physical**	**Mental**	**Physical**	**Mental**	**Physical**	**Mental**
PF	0.75	0.17	0.75	0.03	0.80	0.09	0.82	0.04
RP	0.86	0.19	0.78	0.25	0.80	0.19	0.66	0.36
BP	0.51	0.52	0.53	0.51	0.64	0.28	0.72	0.21
GH	0.37	0.66	0.32	0.66	0.46	0.56	0.71	0.27
VT	0.21	0.88	0.16	0.83	0.16	0.84	0.38	0.60
SF	0.45	0.60	0.48	0.56	0.38	0.61	0.21	0.65
RE	0.69	0.34	0.62	0.36	0.30	0.54	0.08	0.78
MH	0.13	0.89	0.10	0.86	0.02	0.90	0.20	0.78

**Strong ($r \geq 0.70$); Moderate to substantial ($0.30 < r < 0.70$); Weak ($r \leq 0.30$)*